

Where does Russia head now?

Against some odds, President Boris Yeltsin has won the part of the Russian referendum that he says matters most, but the really hard work is only just beginning.

MR Boris Yeltsin, the president of Russia, may have won the battle, but he has not yet won the war he has been fighting for more than half a year with the Congress of People's Deputies. That is how it seemed earlier this week, as people began making educated guesses about the outcome of the weekend's referendum. It seems certain that Yeltsin will have won the trust of a handsome majority of those turning out to vote; consistently, he has said that this is the only question he would take seriously. On the other hand, he is likely to have failed to win a majority of the whole electorate for an early reelection of the Congress. That is bad news, and threatens a continuation of the debilitating battle between the government and the parliament that has dragged on since last November.

Perpetuation of the pantomime of the past few months will serve nobody's purpose. It is true, of course, that for the first time in 75 years, Russia now has in place the mechanisms of national democracy — a government whose president has been elected and a parliament able to enforce, by argument and voting, restraints on what the government does, as well as a constitutional court. But the Congress in its present form dates back to Mikhail Gorbachev's; it was not elected by general suffrage, but (on behalf of the whole of what was still the Soviet Union) partly elected on a local basis and partly nominated by special interest groups. (The All-Union association of philatelists had a separate influence, for example.) The old Central Committee's placemen are still there. No wonder that the government and the Congress have been at loggerheads.

Some of the consequences have been absurd. Early on, the Congress won the right to control Russia's central bank, which has since kept on printing money with which to subsidize bankrupt enterprises, thereby undermining Yeltsin's efforts to engineer economic reform. During the same period, it has voted to change the Russian constitution — itself new — on roughly 50 occasions. Who can be surprised that ordinary people do not know where they stand? And while irrational resentment abounds in Russia (and in the Congress) about the way in which economic reform has created several battalions of dollar-millionaires, many of Yeltsin's problems stem from the sharp practices that have allowed these small fortune to accumulate, mostly unhindered by taxation.

So what should Yeltsin now do? Forbidding though the prospect may be, he needs another talking shop — a commission whose function would be to draw up a definitive Russian constitution that could then be entrenched against the amending whims of the Congress. (Amendment only by referendum

would make sense.) Legislating for seemly relationships between the government and the Congress may be the most urgent need; creating the circumstances in which the patchwork of titularly and hopefully autonomous regions that make up the Russian Federation have an incentive to hang together is, in the long run, more important. Yeltsin could do his country a power of good by directing public attention to these issues.

Can he succeed? There are two views of what has happened in Russia in the past few years. One is that it has lapsed into chaos. Another is that it has survived this far only because of a remarkable social cohesiveness (in which traditional fatalistic passivity may have played an important part). Both views embody aspects of the truth. On the whole, the durability of this remarkable society is the more impressive. Yeltsin has that on his side. But he must do more to suppress the air of lawlessness that now abounds (which would incidentally take wind out of the Congress's sails) and persuade people that paying taxes is a social virtue (which would then help balance Russia's books).

No part of Russia's social fabric is more durable than its network of people with research at heart. To be sure, many able people have now gone elsewhere. Others have been forced out by impoverishment. It is far from clear that the hope of the Russian Academy of Sciences that industrial contracts can keep the rump of its establishments in being can be realized; experience in other countries is not encouraging. Yet there remains a core of Russian research, not simply academic, that remains excellent and which seems certain to survive. Yeltsin could help by ridding it of bureaucratic incubuses. He will not in the long run profit from arrangements that turn the surviving research institutes into training grounds for emigrant specialists.

But would it matter all that much if Yeltsin failed? Russia, after all, is not the only member of the Commonwealth of Independent States. Do not the other eleven deserve at least as much attention, even compassion? And what of the states of Central Europe, themselves also recently unshackled? The answer should be, "Of course". There is indeed a serious danger that Western preoccupation with Russia's plight will induce neglect, even complacency, about the remainder of the ex-Soviet empire. Yet what happens in Russia is distinctively important. Its sheer size is one consideration, its huge stock of military equipment is another. And Yeltsin's failure would probably presage a revival of the old nationalism that drove the tsars to the Pacific coast of Asia. Who would benefit from that? □



Anglo-US concern

Both countries want industry to be more competitive, but differ in their recipes for using science and technology.

Do US president Bill Clinton and British prime minister John Major have more in common than would have been guessed from their brief meeting earlier this month? Each is dusting off schemes for supporting innovative industry. The Clinton administration is said to be looking again at the tangible objectives underlying the federal government's support for research (see page 776), while the British government, in the shape of the Department of Trade and Industry, is seeking a more direct (if a more avuncular than the American) role in fostering industrial competitiveness. What can they hope to achieve?

These are occasions when warnings that civil servants are poor at 'picking winners' are again in order, misanthropic though their utterance may appear. Even those with relatively short memories in either country should be aware of that. In the 1960s, for example, the US administration was captivated by the doctrine of Research Applied to National Needs (or 'Rann'), the most tangible benefit of which was the improvement of building standards for earthquake resistance. (A decade later, the Pentagon's support for the large-scale integration of circuits on silicon surfaces almost surreptitiously created a new industry.) In Britain a little earlier, Mr Harold Wilson's advocacy of "white-hot technology" bequeathed to the country four uneconomic aluminium smelters.

There is a particular danger in the US inclination to regard the \$14 billion a year the federal government now spends on basic research as a package of goal-directed programmes. It has come to seem like that only by accident. Recognizing that, when several agencies have a finger in the pie of, say, biotechnology, it is prudent that there should be a committee to coordinate their spending, the committee's budget then becomes an identifiable object in itself, to be judged against what may be called 'results'. The difficulty, of course, is that agencies' spending covers a diversity of objectives, from the practical to the basic. Direct identification of benefits is difficult, but the substantial advantage to US industry is the small army of people skilled in the techniques of biomolecular manipulation that research produces. If the administration now seeks a coherent shopping-list of objectives, it had better find a way of accounting for these hidden benefits. Otherwise, not merely investigator-led research, but US industry as well, will be the losers.

British ambitions are mercifully less concrete. Mr Michael Hesletine, the Secretary of State for Trade and Industry, seems to have determined on informed exhortation rather than explicit research support. On the face of things, he has luck on his side. There is some evidence that the long recession is ending early in Britain, while the Confederation of British Industry says that companies have begun spending more on innovation even when funds have been squeezed. But, with Britain now ranked eighteenth among members of

the Organization for Economic Cooperation and Development (OECD), he will need all the luck that he can get. □

Too costly software?

Software piracy is wrong, but more flexible policies for selling it would help combat the crime.

STEALING software by illicit copying is wrong, of course, and in many places is a crime. So it should be. Like piracy of the copyright in any kind of work, theft deprives the originators of the rewards to which they are entitled. But the software houses (as, a little pretentiously, they like to be called) are not quite the innocents they pretend. In particular, they differ from those who sell other kinds of copyrighted products, book publishers for example, in their inflexibility over price.

Take, for example, the purchase of a wordprocessing program for a personal computer. In the past decade, the market leaders in the field have been made enormously more sophisticated. Some are almost indistinguishable from desktop publishing systems. So what happens if a purchaser claims that he wishes to acquire only the basic wordprocessing part of the sophisticated package, perhaps because he or she wishes the output to be compatible with a machine on which all the bells and whistles have been legitimately installed? Usually there is no legitimate solution. Buy the whole package, or do without is what the salesman says.

Book publishers, many of whose products are topical, have a solution to this problem: first they sell a hardback edition and then, when the cream of the market has been satisfied, they sell a paperback at a fraction of the cost. The analogy with computer software is not, of course, exact. A computer program (like a dictionary) is a tool that may be repeatedly in use. Even so, the idea that the only versions of the tool available for sale should be more elaborate (and expensive) than many users need is a restraint of trade. In the long run, it may even hurt the owners of the copyright by making illicit copying seem more acceptable.

If software programs were covered by patent rather than copyright legislation, there would be ways round these difficulties. The owner of a patent does not have the unfettered right to make an invention available only on onerous terms. Rather, a patent is an agreement between an inventor and a government that the former shall enjoy the exclusive right of exploitation of an invention for a prescribed period while the community at large shall have access to its benefits. If the inventor of a novel mousetrap then decided that he would offer it for sale only when installed in a purpose-built house for which extra would be charged, most patents courts would probably be prepared to grant compulsory licences to those willing to sell the mousetrap without the house.

Copyright protection is, by contrast, absolute. Even so, it is a question whether the interests of the software houses themselves would be better served if they submitted, perhaps voluntarily, to the unpackaging of their expensive packages (against payment of a royalty). None of this says that software theft should be made legitimate. □

Germany stumbles on enacting plan to integrate eastern scientists

Berlin. The fall of the Berlin Wall also destroyed the regime within which most scientists in East Germany had worked all their lives. All research institutes were closed down, university professors were sacked and a new, smaller research system was created out of the ashes. Now these scientists are being asked to bear another burden: nearly 2,000 who were promised university positions may now find themselves jobless because the universities do not want them.

Nearly all research in communist East Germany was carried out in the 57 institutes of the Academy of Sciences, with universities being given little chance to conduct serious research. A new law established in October 1990 after reunification — the Higher Education Renewal Programme — was intended to mould East German research into a western system based on the unity of teaching and research.

But the transition has not been easy. Academy institutes, which in 1989 employed 24,000 people, were grossly overstaffed. Many employees were engineers who built scientific equipment that they were unable to buy from the West and who became superfluous when the two German economies were combined. By the end of 1991, when the institutes were closed, migration, early retirement and mobility allowances reduced this number to 15,000 — still far too many.

Germany's science council, the Wissenschaftsrat, was called in to evaluate the work of the institutions and recommended that 6,700 staff be retained in newly created institutes (in fact 7,200 are now employed). It also suggested the integration by the end of 1992 of an additional 1,900 into eastern Germany's 10 universities and 400-odd institutes of higher education. But the plan has hit serious problems because the universities are unwilling to offer the scientists contracts.

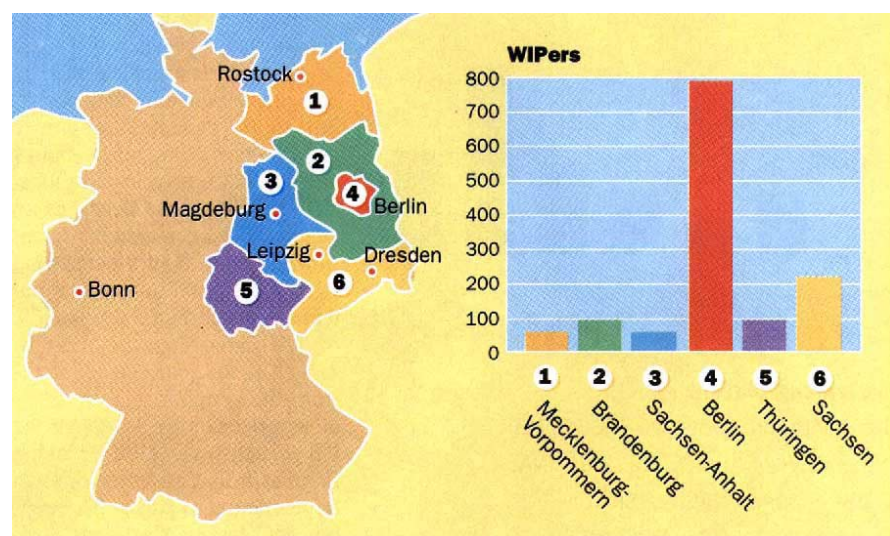
The reintegration programme, known as WIP (*Wissenschaftler-Integrationsprogramm*), was originally given DM400 million (US\$250 million) for four years. Applicants were required to submit a research plan, and successful candidates (now known as WIPers) would be paid for a year of work at one of the new institutes. They would use this time to seek a university job, helped by a coordinating body, KAI, set up for this purpose. Once hired, the scientists would continue to be paid by the federal government, with regional government picking up an increasing share of the salary each year. After three years, the WIPers would be taken on as normal university staff.

But problems soon arose. Universities

forced to reduce staff numbers by half after the science council concluded that they were grossly oversubscribed did not appreciate being asked to take in outsiders at a further cost to their own numbers. At the same time, there was no place for the scientists to work because the universities had done virtually

grant applications and working under the pressure of uncertain funding. So far, only 1,234 of 1,920 have entered into negotiations with universities for permanent positions and only 85 have been given contracts.

Some blame the WIPers for the problems. Last month, for example, the ministers



In addition to transferring researchers from research institutes to universities, the Higher Education Renewal Programme dictates a regional redistribution. Those in the WIP programme must be prepared to distribute themselves more evenly through the eastern *Länder*, and less than half will be able to stay in Berlin.

no research. Debate continues on who should pay the rent for laboratories to be set up if the WIPers are taken on.

The universities are also concerned that in three years they may be obliged to give any available permanent position to a WIPer instead of advertising it for open competition, thereby reducing their ability to attract the best talent. And they worry that they will be asked to make a contribution if the WIPers use up the DM600 million now allocated to the programme.

The programme also puts a strain on the *Länder* (regional governments). Although the exact amount of money allocated to the programme is known, the proportion to be paid by the *Länder*, which are normally financially responsible for universities, is still being negotiated. Two weeks ago, the federal and *Länder* governments agreed to a 75:25 split for 1994 and set individual salaries at 84 per cent of the rate for western Germany. But the terms of funding for the last two years of the programme remain uncertain.

WIPers themselves have found the process difficult. Accustomed to having a job for life, they had no experience of writing

of the *Länder* stated that "some WIPers need to make much more effort to integrate into the universities". But Hartmut Schulz of KAI says that such criticism is unfair. There are many reasons why it has been difficult for researchers to form links with universities, he says, with some not even sure that particular departments would survive investigations into the competence and integrity of the professors.

All of these problems slowed negotiations to such an extent that the federal government was forced to double its initial one-year period of support. But no more extensions are expected, and with most WIPers older than 40, those who do not negotiate a contract during this time fear that they may never again work as scientists.

Despite the problems, there is considerable optimism that most WIPers will, by hook or by crook, have been given contracts by the end of the year. But the idea of true integration — that is, a normal university post — now seems little more than a pipe-dream. For many the WIP programme is likely to be a stressful prelude to permanent unemployment.

Alison Abbott

Clinton team wonders if FCCSET is broken

Washington. The Clinton administration is slowly turning its attention to the \$14 billion that the US government will spend this year on basic research, conducting a review of six multi-billion-dollar interagency initiatives and convening a panel of senior administrators to draw up programmes for future budgets that correspond to the president's domestic policy goals.

pamphlet of a hundred pages or more that included detailed budgets of what each agency planned to spend. This year, in contrast, the White House Office of Science and Technology Policy (OSTP) issued a slim volume containing a four-page summary of each programme.

John Gibbons, the president's science adviser and OSTP director, says that the six FCCSET (pronounced 'fix-it') initiatives, named after the Federal Coordinating Committee on Science, Engineering and Technology that supervises them, are "a dynamic list of topics" that are being reexamined for their contribution to national economic growth. For example, he points out that the four-year-old global change programme began as an attempt to "understand the scientific processes. Now we are trying to shift attention towards its potential impact on the economy."

The search for practical results as well as for good science stems from the administration's desire to use technology to revive the US economy as well as from the fact that the initiatives have grown large enough to attract the attention of Congress. "In the past three years we have provided NSF with \$600 million for the high-performance computing initiative", says Kevin Kelly, an aide to Senator Barbara Mikulski (Democrat, Maryland), who is chair of the appropriations subcommittee that controls the budgets of NSF, NASA, the Environmental Protection Agency and several other agencies.

"What have we bought with that money? I think that NSF would have a hard time answering that question."

The largest FCCSET initiative, some \$4.3 billion for biotechnology, is also the most troubled. Although the initiative includes programmes operated by the agriculture, energy and defence departments as well as by NSF and NASA, more than three-quarters of its budget is controlled by the National Institutes of Health (NIH), which is primarily interested in funding health-related basic research. As a result, efforts to focus attention on environmental issues, for example bioremediation or bioprocessing, have met with limited success. In addition, the administration has requested a 1994 budget for NIH that is essentially flat, leading one NIH official to point out that "if you don't get any more money, then it's superfluous to talk about new initiatives".

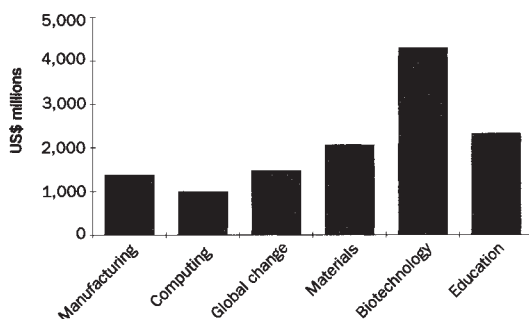
Gibbons's predecessor, D. Allan Bromley, believed that the government could carry out no more than five or six major initiatives at

once because of the time and resources required to coordinate activities among as many as a dozen agencies. Although Gibbons has not said whether he agrees with that analysis, he has talked about "a rotation" in which new programmes replace older initiatives.

One proposal, in the works for the past two years but delayed by the change in administration, would deal with coastal ocean zones. Officials at the National Oceanic and Atmospheric Administration (NOAA) have already done an inventory of existing research programmes and begun to plan an interdisciplinary, global strategy to address issues ranging from non-point sources of pollution to sustainable development. But a meeting last December was canceled by the outgoing Bush administration, and NOAA officials are waiting to hear if the Clinton administration wants to pursue the matter.

That decision — and many others involving the course of science over the next four years — may rest with a new group of

Six research themes in the 1994 budget



The leading players:

Manufacturing: DOD, DOE, DOC, NSF

Computing: DOD, NSF, DOE, NASA

Global change: NASA, NSF

Materials: DOE, DOD, NSF

Biotechnology: HHS, DOE, NSF, USDA

Education: NSF, DOD, HHS, ED

DOD: Department of Defense, **USDA:** US Department of Agriculture, **DOC:** Department of Commerce, **DOE:** Department of Energy, **ED:** Department of Education, **HHS:** Health and Human Services, **NASA:** National Aeronautics and Space Administration, **NSF:** National Science Foundation

Despite the political rhetoric about change, the budget that President Bill Clinton proposed on 8 April for fiscal year 1994, which begins on 1 October 1993, makes only minor alterations in the way in which the government invests in basic research. Even the economic stimulus package that Congress defeated last week, which contained nearly half a billion dollars for the National Science Foundation (NSF), the National Institutes of Standards and Technology and other federal research agencies (see right), would at best have strengthened existing programmes rather than charting new directions.

One sign that the administration is beginning to think seriously about basic research is its decision last week to release only partial information on six interagency initiatives that, taken together, will cost \$12.5 billion. Previously, each initiative — on advanced materials, manufacturing technology, biotechnology, high-performance computing, global climate change and science education — was described in a

Research budgets suffer with defeat of stimulus plan

Washington. The defeat of the economic stimulus package proposed by US President Bill Clinton means that the National Science Foundation (NSF) is unlikely to receive anything close to its 1994 budget request. The stimulus plan contained \$207 million for NSF in the current year, nearly half of the increase of 15 per cent being sought for fiscal year 1994, which begins on 1 October; without that running start, Congress is extremely unlikely to give NSF such a large increase in the face of pressure to reduce the federal deficit and to fund other domestic programmes.

At the same time, two other research programmes that stood to benefit handsomely from the stimulus plan — the \$103 million proposed for the Advanced Technology Program (ATP) within the Department of Commerce and \$47 million sought for cooperative research agreements between industry and the national laboratories in the Department of Energy — are expected to remain important priorities for the Clinton administration and to enjoy rising budgets in 1994 despite the temporary setback in Congress. In the short run, ATP officials say they will cancel plans for another round of applications this summer to their programme, which funds proposals from individual companies and industry consortia, while the Energy Department expects to continue its emphasis on funding research projects within the laboratories that also serve the needs of industry. **J.M.**

about 30 senior administrators convened two weeks ago by Gibbons and Bowman Cutter of the National Economic Council. The so-called deputies' science and technology working group was asked to nominate people within their agencies to serve on panels that will develop policies for 15 topics ranging from specific issues such as medical technology to basic research itself. The goal of the group, according to Tim Newell of OSTP, is to propose specific ways to carry out the president's technology initiative announced in February, but the group is also expected to make recommendations for programmes to be included in the budget for fiscal year 1995 that will be submitted to Congress next winter.

Part of the reason for the uncertainty surrounding the new administration's plans for science are the large number of unfilled positions at OSTP and the ambiguous relationship between it and the office of Vice President Al Gore, the leading figure within the administration on technology issues. Gibbons is believed to be planning a realignment of portfolios, elevating the environment and lowering the life sciences in the OSTP hierarchy, but only one of his choices for the four vacant associate directorships has as yet been cleared by the White House and nominated by the president.

Jeffrey Mervis

Britain backs down over threat to run universities

London. The British government has backed down under pressure from the House of Lords over a controversial proposal that would have given the Secretary of State for Education formal responsibility for the way in which universities are run (see *Nature* 362, 275; 1993).

The government had argued that its proposal to include words expressing the minister's responsibility for further and higher education in a new bill would have had little impact, because the main thrust of the bill in question was changes in the administration of schools. However in a parliamentary debate last week, members of the House of Lords, reflecting concerns expressed by many university vice-chancellors, claimed that the offending clause could be used to justify direct intervention by the government in university administration — and was, as such, an infringement on academic freedom.

Baroness Blatch, minister of state in the Department for Education, announced that the government had agreed to withdraw its proposal and that she would seek discussions with its critics on a more acceptable formulation of the secretary of state's responsibilities.

David Dickson

Choice to head NIST defines shift in US technology policy

San Francisco & Washington. The nomination of Arati Prabhakar as director of the US National Institutes of Standards and Technology (NIST) clearly shows the contrast between the views of President Bill Clinton and former President George Bush on industrial policy.

Three years ago, Craig Fields was forced out as director of the Defense Advanced Research Projects Agency (DARPA) after Prabhakar, then deputy director of DARPA's defence sciences office, made a \$4 million investment in Gazelle Microcircuits Inc., a tiny Silicon Valley semiconductor company that makes high-speed gallium arsenide chips. The agreement angered the Bush administration, which was philosophically opposed to the government becoming a venture capitalist for fledgling companies. Although the government eventually decided to accept royalties from sales of the chip rather than to take an equity share in the company, it was widely seen as the last straw in a conflict between Fields and the administration.

Now Prabhakar, a 34-year-old applied physicist, is poised to become head of NIST, to which the Clinton administration has given a leading role in stimulating the US economy. Not surprisingly, her former boss is delighted by her promotion.

"She is a very, very capable manager", says Fields, chief executive officer of the industry-led Microelectronics Computer Consortium in Austin, Texas. "She's innovative and she understands new technology and how to apply it."

Although Gazelle's technology was impressive, the company was unable to survive on its own. In 1991, a year after the Pentagon's investment, Gazelle merged with two other chip makers, TriQuint Semiconductor of Beaverton, Oregon, and GigaBit Logic of Newbury Park, California. TriQuint had been making the chips designed by Gazelle's researchers, and its name was retained by the new company, which last year had a turnover of about \$30 million and showed a small profit.

The US government continues to receive royalties from the sale of chips designed under the original agreement with Gazelle, although TriQuint is no longer receiving any support from ARPA (the 'D' was dropped last month as part of the Clinton administration's promise to shift spending from the military to the civilian sectors). Spencer Brown, TriQuint's executive vice president and chief financial officer, says that the federal contribution enabled Gazelle to continue improving its technology and that the company expects increased demand for its products if a national information superhighway becomes a reality

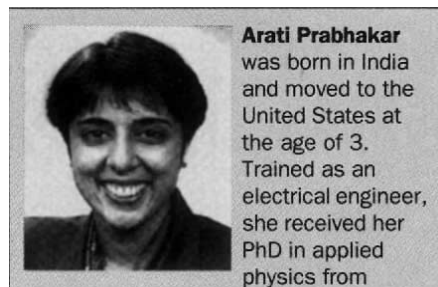
(see *Nature* 362, 582; 1993).

Fields is confident that Prabhakar, who now directs ARPA's \$300-million microelectronics technology office, can run an agency whose budget is expected to quadruple to \$1.4 billion in the next four years. But he warns his former coworker that her every move will be scrutinized.

"If one tries to help every industry, you spread yourself too thin", says Fields. "But if you favor those with the best chance of succeeding, then you leave yourself open to political attacks from those left out."

The Clinton administration hopes to prevent that from happening. Speaking last week at a conference entitled "New Directions in Technology Policy", Commerce secretary Ron Brown said that "a raging debate about the proper relationship between industry and the government" is over and that the government is "ready to work hand-in-hand with industry to strengthen the US economy". If nothing else, Prabhakar's appointment seems to be proof of that commitment.

Frederic Golden & Jeffrey Mervis



Arati Prabhakar was born in India and moved to the United States at the age of 3. Trained as an electrical engineer, she received her PhD in applied physics from

California Institute of Technology in Pasadena and moved to Washington in 1984 to become a congressional fellow at the Office of Technology Assessment.

In 1986 she joined the Defense Advanced Research Projects Agency (DARPA), first as programme manager of the electronics sciences division of the Defense Sciences Office and, later as deputy director of the office, where she negotiated the agreement with Gazelle (see left). Two years ago, she was asked to head a new office on microelectronics technology, with an annual budget of \$300 million.

Although Prabhakar has declined to talk to the media until after she is confirmed by the US Senate, three years ago she described her job at DARPA as one of "making sure that things happen. My goal is to get the technology out of the lab and into the hands of others." That statement is also an apt description of what she will be asked to do at NIST.

F.G. & J.M.

US government and utilities revive California solar project

Washington. The US Department of Energy this week gave the go-ahead for a \$39-million project to rebuild the core of one of the world's largest solar power plants with the goal of generating commercial quantities of electricity by 1998.



Going with the flow in California

The Solar Two project at Barstow, California, is a successor to a 10 MW power plant that ceased operations in 1988. Combined with the sharp increase in solar energy research proposed by US President Bill Clinton for fiscal year 1994, which begins on 1 October, the new agreement between the department and an industrial consortium is

a powerful symbol of revived federal interest in renewable energy research.

The original demonstrator at Barstow used arrays of mirrors to focus the sun's rays on a central 'power tower' containing steam that was heated directly to generate electricity. This approach was abandoned after five years because the tower's output was erratic, changing when clouds passed overhead, for example.

Now the moth-balled site is to be revived with extra mirrors — or heliostats, as solar researchers call them — and an energy cycle that will heat molten nitrate salt to 565° C and store it to generate a steady flow of steam for the existing turbo-generators.

The use of the molten salt cycle has been proved on a smaller scale at the department's Sandia National Laboratory at Albuquerque, New Mexico, which has already spent \$30 million. Significant technologies under development there includes improved stretched-membrane heliostats and pumps and valves to handle the molten salt.

The department will provide \$19.5 mil-

lion over the next three years to build Solar Two, a sum to be matched by the South California Edison Company and ten other industrial partners. Construction will begin in the summer of 1994 for operations in 1996 intended to prove the technology can be used in commercial 100 MW solar power stations.

A proposed budget increase of \$70 million for renewable energy, to a total of \$327 million, is spread evenly across photovoltaic research, thermal solar energy work (including Solar Two), biofuels and wind power. The department also intends to spend \$16 million to find commercial applications for existing renewable technologies.

The strongest opposition is likely to come from supporters of nuclear power, which was cut sharply in the Clinton budget. But advocates of renewable energy point out that the new budget still leaves the field far behind where it stood more than a decade ago, when President Jimmy Carter spent as much as \$800 million a year on renewable research.

Colin Macilwain

Framework budget finds its level

Strasbourg. The council of research ministers for the European Communities (EC) today (29 April) will discuss a proposed budget for the fourth Framework programme that matches what member countries have said they can afford.

The budget proposed last week by EC research minister Antonio Ruberti for Europe-wide research requests ECU13.1 billion (US\$15.8 billion) over the next four years. It is the maximum acceptable level, at 1993 prices, agreed to by the council of ministers and, if accepted, would represent 4 per cent of the EC's total budget. Ruberti's predecessor, Filippo Pandolfi, had asked for ECU14.7 billion, which the ministers rejected as too expensive at their last meeting in December. The current Framework programme has a budget of ECU6.6 billion.

The expansion of the programme reflects the EC's commitment to research and to the mobility of research personnel. It includes for the first time a modest amount of money for social science research, and emphasizes closer ties to international research laboratories such as CERN, the European Laboratory for Particle Physics in Geneva, and EMBL, the European Molecular Biology

Laboratory in Heidelberg, as well as closer cooperation with national research organizations.

In particular, the Framework budget is divided into seven categories: information and communications technologies, 36 per cent; energy, 23 per cent; industrial technology, 16.5 per cent; life sciences, 12 per cent; environmental research, 9 per cent; a European transportation policy, 2.5 per cent; and socio-economic research, 1 per cent.

The European Commission, which is responsible for formulating the proposal in line with the views of the council of ministers, was criticized for the repeated delays in the adoption of the third Framework programme and hopes to stick to the intricate timetable for approval of the new programme, which must be accepted at three levels. The council of research ministers is expected to reach a decision by the beginning of June. At the same time, the European Parliament is considering the document and expected to finish its work by the beginning of July. The goal is a political agreement in July between the council, commission and parliament — but acceptance rests with ratification this summer of the Maastricht Treaty.

Alison Abbott

NEWS IN BRIEF

Washington. The US National Aeronautics and Space Administration (NASA) has expanded the length and scope of its mission in December to cope with the growing list of faults on the Hubble Space Telescope. The new schedule calls for five spacewalks rather than the three originally planned for a mission lasting 11 days rather than eight. Two pairs of astronauts from the mission's seven-member crew will alternately perform the spacewalks, a record number that may be further increased to seven. NASA says that the extension of the mission would not add substantially to the cost.

The repairs to be carried out include the replacement of gyros, solar arrays and cameras and the installation of a device to correct the optics of the troubled telescope, launched in April 1990.

C.M.

Paris. France's Genetic Engineering Commission last week approved as safe the country's first gene therapy trial for cystic fibrosis. The decision opens the way for the trial to begin this autumn pending approval from the French drug agency and the local ethics committee. The trial, submitted by Transgène, a French biotechnology company, and supported by the French Cystic Fibrosis Association, will test the feasibility and tolerance of aerosol administration of the CFTR gene. Gabriel Bellon of Lyon will carry out the trials on several adult volunteers. The CF-modified adenovirus vector, developed by Transgène and Michel Perricaudet of the Centre National de la Recherche Scientifique viral genetics laboratory at Villejuif, is the same as that used earlier this month by Ronald Crystal of Cornell University in New York on a 23-year-old patient in the first CF gene therapy trial.

D.B.

Clinton announces package of environmental reforms

Washington. US President Bill Clinton used last week's celebration of Earth Day to announce several initiatives that, despite their lack of detail, represent his strongest actions as yet on environmental protection.

As expected, Clinton said he would sign the international biodiversity treaty endorsed

Quality, which Clinton wants to disband, of enforcing the National Environmental Protection Act and mediating disputes between federal agencies over environmental issues. The Senate may vote this week on a bill to give the proposed department that authority, but the issue is expected to remain controversial when it moves to the House of Representative.

Cliff Owen, UPI

So far the EPA has taken a back seat to the Department of the Interior in shaping new environmental policy in the new administration; Interior Secretary Bruce Babbitt has already taken action on endangered species, land management issues and reorganizing its science programme. In his Earth Day speech, Clinton officially announced the creation of a new national biological survey within the Interior Department (see *Nature* **361**, 574; 1993).

The survey will have an annual budget of \$180 million, with more than \$100 million coming from the US Fish and Wildlife

Service, and draw on 1,600 employees seconded from within the department. Its duties will include an inventory and monitoring of biological resources as well as providing independent scientific advice on ecological science to officials throughout the government. The scientists reassigned to the survey possess expertise in areas such as population dynamics, physiology, animal behavior, habitats and biodiversity. The reorganization is expected to go into effect on 1 October pending congressional approval of the internal transfer of funds.

In the next few months, the Clinton administration expects to conduct a broad review of federally sponsored environmental science and research. One intent of the review, says Nancy Maynard, an assistant director at the White House Office of Science and Technology Policy, is to eliminate duplication but it is not likely to lead to a wholesale restructuring of agency missions.

A proposal for a National Institute for the Environment to coordinate and fund scientific research on the environment, modelled on the National Institutes of Health, may soon gain the support of the National Academy of Sciences. Although some believe that the institute is unnecessary, an academy panel is expected next month to endorse the basic concept and to propose a possible structure. Representative Jim Saxton (Republican, New Jersey) is preparing legislation to create the institute.

Tony Reichhardt

US-German council formed to strengthen international ties

Munich. Germany and the United States have formed "an intellectual bridge across the Atlantic" to shore up weakening links between scientists and scholars in the two countries.

The German-American Academic Council was set up this month on the initiative of Germany to support projects and exchanges in science, including social science, and technology. The council will be sponsored by three major research organizations in Germany — the Deutsche Forschungsgemeinschaft (DFG), the Max Planck Society and the Alexander von Humboldt Foundation — and in the United States by the National Academy of Sciences, the American Council of Learned Societies, the American Academy of Arts and Sciences and the Social Sciences Research Council.

So far, Germany is footing the entire bill. The German Ministry for Research and Technology has committed DM18.1 million (US\$11.5 million) over the next four years. US sponsors are being sought among the various private research foundations, but according to Reimar Lüst, president of the Humboldt Foundation, "there is no hurry, the council has enough money to get started".

The idea, conceived by Chancellor Helmut Kohl and endorsed last month in talks between the German leader and US president Bill Clinton, is an attempt to revive the strong transatlantic ties that existed immediately after the Second World War but which have steadily declined as that post-war generation ages. Kohl hopes that the council will become a way to revive these links at a time when Europe is very much looking inwards.

Lüst sees the council as "an intellectual bridge across the Atlantic . . . that can be an important vehicle for joint policy studies and interdisciplinary cooperation" in such areas as industrial policy, technical education and the impact of immigration on the national economy. A US-German committee that issued a preparatory report on the need for such a council sees it as a forerunner for an institution to promote cooperation between the United States and all of Europe.

The council will be composed of 30 members, 15 from each country, whose election should be complete within the next few weeks. A main office will be set up in Bonn with a branch office in Washington. The first council meeting is planned for early summer, when the first project grants, for up to 18 months, are expected to be made. Applications from younger scientists will be given priority, as will projects in the humanities.

Alison Abbott

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Clinton turns green on Earth Day.

last summer by more than 160 nations in Rio de Janeiro (see *Nature* **362**, 577; 1993). And he took the side of Vice President Al Gore against other cabinet members in committing the United States to lowering by 2000 its emission of greenhouse gases to 1990 levels, a provision the Bush administration had kept out of the global climate treaty signed at Rio. How this will happen — whether, for example, the US programme will require tough new regulations or rely on incentives — is still to be determined; Clinton has asked for a "cost-effective" plan by August.

Clinton also announced measures to clean up the government's own house, including a commitment that federal agencies will buy ozone-friendly products, energy-conserving computers, fuel-efficient vehicles and recycled products. Clinton ordered agencies with facilities that release toxic pollutants, including military bases and research laboratories, to develop a plan to reduce their output by half by 1999, and to report them to the public as private companies are now required to do.

In his speech at the National Botanical Gardens in Washington, Clinton made only passing reference to efforts to elevate the Environmental Protection Agency (EPA) to a cabinet-level Department of the Environment, saying that he hoped it would happen "soon, by the grace of Congress". Environmental groups are concerned that the new department will be less capable than the White House Council on Environmental

Canadian study fuels move to limit rise in drug prices

Quebec. The Canadian government body charged with holding down the cost of patented drugs is proposing amendments to its pricing guidelines at the same time a study shows that Canadian prices are often higher than the international median. The Canadian parliament recently extended patent protection after pharmaceutical companies promised to control prices and to increase their spending on research into new drugs.

A new study by the Patented Medicine Prices Review Board (PMPRB) found that 56 per cent of the introductory prices of 124 new drugs in Canada were greater than the corresponding median international prices. And the most popular drugs were more expensive in Canada than in any other country listed in the Patented Medicines Regulations except the United States and Germany. The Canadian analysis confirmed the findings of a recent US General Accounting Office study that found prices for the same drug to be higher in the United States than in Canada.

The study found that the less useful the new product, the more the Canadian drug companies charged for it. Prices were higher in 21 per cent of cases involving drugs that provide moderate, little or no improvement over comparable medicines, but in only 9 per cent of cases involving therapeutic breakthroughs or drugs providing substantial improvement. Among products providing new strengths or a comparable dosage of an existing medicine, only 5 per cent had higher Canadian prices.

The board found that 105 of 177 drug products sold in Canada were priced above

the median international price. In 42 cases the Canadian price was the highest in the world.

Patent drug prices have been under intense scrutiny in Canada since 1987 when parliament approved an extension to the patent act that critics warned would lead to higher drug prices. Brand-name pharmaceutical companies denied that this would happen and promised greater investments in research in return for longer patent protection.

The price review board was set up as a result of the new patent legislation. It has found that between 1987 and 1991, the price of existing patented drugs increased by 2.9 per cent compared with an increase of 4.7 per cent permitted by the guidelines, and that the annual increase has consistently remained below the rate of inflation.

In return for the 1987 extension, patent drug companies promised to raise their research spending over the next ten years from 4.9 per cent of sales to 10 per cent. In October 1991, the Pharmaceutical Manufacturers' Association of Canada (PMAC) said that 47 of its 67 members were spending 10.1 per cent of their sales on research, a total of \$301 million.

Last month, the government passed legislation extending market exclusivity for most patent drugs for another three years and modifying Canada's system of 'compulsory licensing', which makes it relatively easy for generic drug companies to copy brand-name products (see *Nature* 359, 351; 1992). In lobbying for the law, the brand-name manufacturers announced that their research spending exceeded \$400 million a year.

The revelation that Canadian prices for new drugs are relatively high internationally has become a political issue. "Something is wrong here", says New Democratic Party leader Audrey McLaughlin. But PMAC's financial officer, Robert Livingstone, says that drug prices depend on many factors, including the type of national health system through which they are marketed, and that meaningful comparisons are difficult. "We have not yet been able to review the data", he says. "We are committed to fair and reasonable prices, but what is fair and reasonable?"

Last autumn, the price board proposed amendments that, among other things, would limit the price of a new drug that provides moderate or no improvement over existing drugs to the median international price or the price of all drugs in that category. Hearings on the proposals were held in February, and the board is expected to make a final decision in June.

David Spurgeon

15 drug companies to share data on AIDS therapies

Washington. In a break with tradition, 15 pharmaceutical companies from the United States and Europe last week announced an agreement to work together on combination drug therapies to fight the human immunodeficiency virus (HIV). The participating companies will share information and supplies of drugs early in the drug development process to carry out independent testing of promising combinations of new drugs. The collaboration, which includes an effort to standardize preclinical testing procedures, does not extend to the search for an AIDS vaccine nor to basic research on the disease.

A recent survey by the US Pharmaceutical Manufacturers Association reported that more than 90 medicines were being developed to combat AIDS and AIDS-related conditions. But progress has been hindered by problems with toxicity and with the ability of HIV to become resistant to individual drugs. Earlier this month, preliminary results from an Anglo-French trial cast doubt on the ability of AZT (azidothymidine) to delay the onset of AIDS in patients with HIV (see *Nature* 362, 483; 1993).

A year in the making, the agreement was conceived by P. Roy Vagelos, chairman and chief executive officer of Merck, and Edward Scolnick, president of Merck Research Laboratories, with strong support from Juergen Drews of Hoffmann-La Roche. Participants in the Inter-Company Collaboration for AIDS Drug Development will share relevant information and supplies when a drug reaches phase II clinical testing, the point at which the drug's safety profile is known and preliminary tests are under way on efficacy.

Participants must be actively involved in AIDS antiviral research and have an investigational new drug in the early stages of drug development. By dealing with work done before a product is ready for market, the collaboration hopes to avoid violating US antitrust laws and other problems associated with intellectual property rights and the sharing of royalties.

Although most of the important companies have signed up, two notable absentees are Abbott Laboratories and Upjohn. A spokesperson for Upjohn says that the company prefers to maintain an independent anti-AIDS programme, which now receives more than a fifth of the corporate budget for drug discovery research.

The participating companies, nine of which are from the United States, include AB Astra, Boehringer Ingelheim, Bristol-Myers Squibb, Burroughs Wellcome, Du Pont-Merck, Eli Lilly, Glaxo, Hoechst AG, Hoffmann-La Roche, Merck, Miles, Pfizer, Sigma Tau, SmithKline Beecham and Syntex.

Diane Gershon

Zeneca increases research spending

London. The various business that will make up Zeneca, the new company being created out of the biosciences activities of Britain's Imperial Chemical Industries (ICI), spent £457 million (US \$640 million) on research and development last year, according to figures released last week in preparation for the company's stock market launch in June.

This figure represents 11 per cent of the company's total sales and is an increase of 9.8 per cent over R&D spending in 1991. In contrast, growth in R&D spending between 1990 and 1991 was only 5 per cent, largely as a result of a rationalization of research on agrochemicals. Peter Doyle, formerly research and technology director of ICI who is to become an executive director of Zeneca, estimates that 1993 spending is likely to rise to "just under £500 million".

D.D.

Japanese computer makers enlist scientists overseas

Tokyo. Japanese computer makers are turning to scientists overseas to help them develop software and applications for massively parallel supercomputers. Next month, NEC will ship a Cenju-2 parallel computer to the new Swiss Scientific Computing Center (CSCS) in Manno, while Fujitsu already has links with Australian scientists and others around the world.

The US computer industry has long used academics to develop software for new high-performance computers. But such links are rare in Japan, and the interaction with foreign scientists is also a recent phenomenon.

In 1991, Fujitsu loaned at no cost a prototype massively parallel computer, the AP1000, to the Centre for Information Science Research at the Australian National University (ANU). The centre, Australia's leading supercomputer facility, in 1989 began a five-year, \$10-million agreement with Fujitsu for joint development of supercomputer software. The company also has three AP1000 computers at its laboratory in Kawasaki that serve users in Switzerland, Britain and the United States.

Fujitsu says that the number of foreign users has tripled in recent months and that 21 groups outside Japan are linked to the computers. There are also about 30 visiting foreign scientists at Kawasaki using the machines. Outside scientists are developing software for generic problems in molecular dynamics, quantum physics, fluid dynamics and high-energy physics as well as parallel software algorithms, computer graphics and circuit design and device simulation of more immediate use to Fujitsu.

Julian Clarke of the University of Man-

chester's Institute of Science and Technology in Britain uses on-line connections via Internet "several times a day" to develop molecular level simulations of polymers and other amorphous materials. "We are able to perform scientific calculations that otherwise would not be feasible", he says. In return, Clarke provides Fujitsu with the computational chemistry software.

The new NEC High Performance Computing Software Development Center in Switzerland, which opened last month in the same building as CSCS, will by the end of this year give a limited number of outside users access to its Cenju-2 parallel supercomputer, says A. Scheidegger, director of CSCS. The centre was established last autumn with SFr40 million (US \$30 million) from the Swiss government and has an NEC SX-3 vector-processing supercomputer linked to the high-capacity Swiss education and research network (SWITCH) and to those carrying out joint research with Swiss government organizations.

Japanese companies have chosen Australia and Switzerland because they lack a domestic industry that might protest against the import of 'free' high-performance computers. Furthermore, both countries have a good track record in software and have extensive computer networks.

The linkage with academics around the world is not confined to Japanese companies. Cray Research Inc. in Minnesota, which has traditionally asked university scientists in the United States to help it to develop software, hopes to work with European scientists to develop specific software.

David Swinbanks

UK companies report increased R&D spending in 1992

London. Spending by British companies on research and development (R&D), which dropped sharply between 1989 and 1991 as a result of the economic recession, appears to be rebounding even before any overall recovery, according to a survey released last week by the Confederation of British Industry (CBI).

The growth has been particularly strong in traditional manufacturing regions, such as the northwest of England and the East Midlands, as well as the high-technology areas of Scotland and East Anglia. It has been lowest in the southeast and southwest of the country, where wealth is based more heavily on service industries.

"The 1992 figures are the highest since the survey began in 1989, and are in marked contrast to 1991's figures, which were the lowest", says Bryan Smith of ICI engineering, and chairman of the CBI's research and innovation committee.

The CBI's conclusions are based on replies from 268 manufacturing companies and 151 non-manufacturing companies to a questionnaire completed at the end of last year. These showed that half of the manufacturing companies increased their R&D spending in 1992 — and that almost as many intended to continue this trend in 1993.

In particular, CBI compared the number of companies increasing (or planning to increase) their R&D spending with those reducing it. Earlier studies by the confederation's NatWest technology unit showed that, between 1989 and 1991, the excess of those increasing R&D spending on new products fell from 40 to 16 per cent. The downward trend was confirmed by figures released in January by the Central Statistical Office showing a 4 per cent drop in business spending on R&D between 1990 and 1991.

In sharp contrast, the returns for 1992 showed an excess of 42 per cent of companies were spending more rather than less on R&D — 9 per cent higher than had been anticipated on the basis of previous declaration by companies of their planned expenditure for the year. (Figures for R&D on new processes showed a smaller increase, from a balance of 15 to 18 per cent; among non-manufacturing companies, the balance of companies spending more on product R&D increased from 10 to 23 per cent.)

The fastest rate of growth in R&D spending was reported by manufacturing companies with between 500 and 1,000 employees, 67 per cent of which reported an increase last year over 1991. In terms of regional differences, East Anglia reported the highest proportion of companies that had raised R&D spending, and the southwest of England the lowest.

David Dickson

Simplify science? I'll drink to that

London. Can you describe what the Higgs boson is — and why you would like to find it — on a single sheet of paper? If so, William Waldegrave, Britain's cabinet minister for science, wants to hear from you.

Waldegrave threw out the challenge to the physics community last week at the annual conference of the Institute of Physics in Brighton. Scientists, he said, must explain to the public what they are doing if they want to receive continued support. More immediately, the winner

will receive a bottle of vintage champagne.

Waldegrave said that the need to promote the public understanding of science would be addressed in the forthcoming White Paper (policy document) on science and technology and that he planned to launch a campaign to "evangelize" science. Hence his proposal for a competition to describe the significance of the Higgs boson — a target of the next generation of accelerators such as the proposed Large Hadron Collider at CERN in Geneva and Superconducting Super Collider under construction in Texas.

"If you can make me understand that, I stand a better chance of helping to get you the money to find it," Waldegrave said. Entries should be sent to him by 1 June at the Office of Public Service and Science, Cabinet Office, 70 Whitehall, London SW1A 2AS.

D.D.



Degrees of national wealth

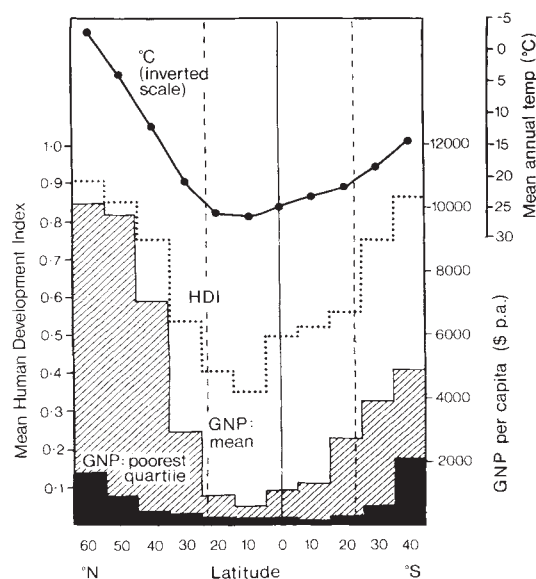
SIR — 'The South' is now widely used as the collective term for non-industrialized nations, with the implication that there is either a consistent trend or a major discontinuity in the relationship between latitude and socio-economic development. Within the Northern Hemisphere both these effects occur, with a marked decline in per capita gross national product (GNP) between high and low latitudes. National data for the United Nations Development Programme Human Development Index¹ (HDI; based on life expectancy and adult literacy, as well as income) behave similarly. However, a fully global analysis of these development indices shows that the geographic distribution of high living standards is bimodal, with a closer association to temperature than latitude (see figure).

Similar patterns in the global distribution of per capita GNP have been found in previous analyses, of 1972² and 1988³ data. The latter study was based on GNP data for 141 countries and territories, ascribed to latitude bands according to the location of capital cities. Because all countries were given equal weighting, regardless of population size, the importance of small, high-income nations (for example the oil-producing Gulf states) was accentuated. In the current analysis, mean per capita GNP was determined from totals of GNP and population for 10° latitudinal intervals, assuming even distribution of 1991 national data (compiled from OECD, World Bank and national sources)⁴ across the latitudinal range individually occupied by 200 countries and territories. Mean HDI values were determined similarly, from 1990 data¹ for 150 countries and territories. The HDI data excluded several non-UN countries, and gave a single value for the ex-Soviet Union (for GNP, separate data were used for the latter).

Not surprisingly, the global distribution of wealth is strongly influenced by the mid- to high-latitude location of the United States and Canada, Japan, the nations of Western Europe, and, in the Southern Hemisphere, Australia and New Zealand. Nevertheless, the same pattern occurs when these countries are excluded; for example, when the analysis is limited to the poorest quartile of countries in each latitudinal interval.

For the 11 latitude bands considered, per capital GNP is more closely correlated with mean annual temperature (r^2

= 0.97) than with latitude *per se* (r^2 = 0.87). Regression analysis for the former relationship indicates a decrease of around \$400 per capita GNP for a 1°C increase in temperature. However, such statistics are descriptive, not predictive, and must be treated with caution. Two provisos are particularly important. First, that non-climatic influences (including the distribution of natural resources) are likely to dominate at the regional level, as evidenced by the 10–200 fold variation in national per capita GNP values within 10° latitude bands. Second, the occurrence of a strong global correlation between GNP and temperature does not necessarily mean that a cooler, more seasonal climate inherently favours socio-economic development: it is arguable that such a relationship could be a fortuitous consequence of the high-latitude origins of modern industrialization, with that situation subsequently



Latitudinal distribution of 1991 (GNP) per capita (US\$ per annum), 1990 Human Development Index (HDI) and mean annual land surface temperature. Land areas north of 65°N and south of 45°S are excluded from GNP and HDI analyses (that is, assumed to be uninhabited); temperature data also excludes land over 1,500 metres.

maintained by political and financial structures that have generally been disadvantageous to tropical nations. An additional consideration is that the latitudinal pattern of agricultural productivity is closely similar to that for per capita GNP. (R. W. Kates & Jen-hu Chang, personal communications), albeit crop yields in temperate regions are enhanced by 'rich country' inputs of fertilizer, pesticides, mechanization and advanced irrigation technology.

Resolution of such issues requires much more detailed demographic and

environmental data^{3,5}, to elucidate the dynamic processes whereby climate interacts with national wealth and human development. In this context, the current North–South terminology may be politically convenient; however, such a radical re-definition of global geography has dubious scientific merit.

Phillip Williamson

Richard Moss

Wenner-Gren Centre,
Sveavägen 166,
S-133 46 Stockholm, Sweden

1. UNDP Human Development Report 1992 (Oxford University Press, 1992).
2. Campbell, D.J. & Katz, C. *Atlas of the Least Developed Nations: Groups and Characteristics* (Clark University, Worcester, Massachusetts, 1975).
3. Kates, R.W. & Haarmann, V. *Environment* **34**, 4–11: 25–28 (1992).
4. Cambessédès, O. *Atlaséco 1993: Atlas Economique Mondial* (Editions du Sérail, Paris, 1993).
5. Clarke, J.I. & Rhind, D.W. (eds) *Population Data and Global Environment Change* (HDP Report 3, ISSC/ UNESCO, 1992).

Population control

SIR — In my experience and that of others, many women in societies with very high birth rates would like to start reproducing later, stop earlier and space their births at wider intervals. If contraceptive services were available locally, cheaply, safely and privately, then many women would use them, even in cultures where most men are indifferent or opposed to contraception and religious teaching is generally against it. It cannot be useful to talk about population growth (*Nature* **362**, 379; 1993) as a "threat to the global commons" that may need to be "dealt with by coercion" before it becomes bothersome to "the rest of us", which I assume means those of us lucky enough to live in countries where contraceptives are readily available. This analysis is particularly misleading because "the rest of us" use up more global commons, both per capita and in total, than any of the poverty-stricken countries where birth rates are high, in some cases by orders of magnitude. If we want to save the fish in the sea, the trees in the rain forests and the ozone layer, the wealthy countries have to look to their own consumption habits. Rapid population growth is a problem, because it increases poverty and ill-health in societies where it occurs. We will have more success at slowing world population growth if we see it as problem of enabling women in poor countries to have control over their own reproduction, which some carefully targeted development aid could go a long way towards achieving.

Ruth Mace

Department of Anthropology,
University College London,
Gower Street, London WC1E 6BT, UK

Forty years of molecular information

Nicholas Short

At UNESCO's celebration of the double helix's fortieth anniversary in Paris last week, nostalgia was overlaid by vigorous optimism about the future of molecular biology.

"WHAT", asked François Jacob (Institut Pasteur, Paris), opening the meeting in Paris on "40 years of molecular genetics" at UNESCO's headquarters on 21–23 April, "will future historians remember of this century?" Various wars, the rise and fall of fascism and communism, and the end of colonialism, perhaps, but surely also the fantastic development of science and industry. In the first half of the century the focus was on physics, but in the second, the emphasis had shifted to biology, triggered in large part by the discovery that Watson and Crick published forty years ago.

The focus of the meeting, one of several being held this year (see, for example, *Nature* 362, 105; 1993) was firmly on the intellectual background from which the discovery sprang. First, Max Perutz (MRC, Cambridge) described the pioneering work of Oswald Avery, proving that genes are made from DNA. Then Francis Crick (Salk) recalled how Avery's work, buttressed by that of Hodgekiss and of Hershey and Chase, had led him to muse on how to copy a three-dimensional structure: as with a sculpture, perhaps, by making a mould? Or, more plausibly, by copying some linear array which had an intrinsic capacity to fold itself up?

Clearly, the structure of the molecule in which the cell stores information was crucial. X-ray-fibre diffraction data had been obtained for both the *A* and *B* forms of DNA by Rosalind Franklin and Maurice Wilkins at King's College, London, but they lacked the phase data necessary to solve the structure. So, unlike Franklin, Crick and James Watson (Cold Spring Harbor) elected to build models and calculate the diffraction pattern each would produce. It was only after they had worked out the correct tautomeric form of each base and found that adenine is exactly the right shape to pair with thymine (and guanine with cytosine) that they thought of Erwin Chargaff's discovery that DNA always contained equimolar amounts of the bases in each of these pairs.

From then, the path to the double helix was clear. Leslie Orgel (Salk) recalled that only two weeks after Watson and Crick's first paper appeared, the entire biochemistry department at Oxford travelled to Cambridge to view the structure and followed this up with a telegram: CONGRATULATIONS! (signed) GENE.

That the two paired strands formed a double helix was in Crick's original view a nuisance; they had then to be unwound.

The key point was that the bases paired, making it possible for the first time to see how a chemical entity might be able to store and copy information. That was the theme that informed the entire meeting.

Of course, it was not always like that. Before the first two helices had ever intertwined, biological information was very probably being copied and stored as RNA. Indeed, that molecule is more versatile than DNA and, as Tom Cech (HHMI, Boulder, Colorado) demonstrated, can form enzymes of an efficiency and specificity as impressive as those of proteins. But whether the storehouse of information is DNA or RNA, the crucial question is the error rate on copying. For any biological text, explained Manfred Eigen (MPI, Göttingen), there is a single error threshold; above this, the error rate is too high, and information is lost on each round of replication, whereas below it, the error rate is too low, and no new variants can ever emerge. By mapping the changes during the evolution of viruses, he has shown that the threshold varies in different codon positions and in different genes.

On a similar theme, the conference organizer, Giorgio Bernardi (Institut Jacques Monod, Paris), brooded on the significance of the many gene-rich segments of the genomes of mammals and birds termed isochores, whose G + C content has been highly conserved. The exact fraction of G + C in these regions may have taxonomic importance, as the values in birds, mammals and cold-blooded vertebrates are distinct.

Information flow of a different kind was explored by Walter Gilbert (Harvard), asking whether genes or their introns came first. Now that the full complement of introns in the ancestral triose phosphate isomerase gene has been mapped, he has been able to show that they are placed so that amino acids near each other in the protein's structure are encoded wherever possible by the same exon, bolstering the idea that the original form of the gene contained a full complement of introns and was produced by exon shuffling.

But the most spectacular example of biological information on the move stemmed from Takashi Gojobori (National Institute of Genetics, Mishima)'s project to map the MHC haplotypes of all the races of humanity. Not only are many of the possible haplotypes confined to one race or area, but there is even a haplotype that is confined to the former 'Silk Road' be-

tween China and Europe.

If DNA is the repository of biological information, access to that information is controlled and directed by proteins. In a sweeping review of the various domains by which proteins bind to DNA, Aaron Klug (MRC, Cambridge) revealed the diversity of structures that have evolved to meet the requirements of physical stability and evolutionary flexibility imposed by the need to recognize only a tiny fraction of the binding sites available in an entire genome.

Although most DNA is in the *B* form, whose structure was originally determined by Watson and Crick, there is no reason why the cell should not make use of other forms as well, and indeed Alex Rich (MIT) described a protein binding with high affinity to the left-handed helix of Z-DNA that he first described. Most of this form of DNA is confined to actively transcribed genes, where Rich speculates that it may function to keep an appropriate distance between successive molecules of RNA polymerase.

Even without this complication, however, the need to allow the transcriptional machinery access to the DNA without permanently disrupting its protein packaging is a demanding one. For many years, Andrei Mirzabekov (Engelhardt Institute, Moscow) has been probing the way in which vertebrate cells perform this feat using an increasingly sophisticated series of techniques for mapping protein-DNA contacts *in vivo*. He finds that although the globular bodies of the histone molecules no longer contact the DNA of heavily transcribed genes, their charged tails still do, leading him to suggest that they remain tethered by their tails while the polymerase passes so that they can rapidly 'snap back' into their former structure.

Approaching this problem from the opposite direction, Gary Felsenfeld (NIH) described a series of experiments mapping the redistribution of nucleosomes caused by the passage of T7 RNA polymerase *in vitro*. Under these circumstances also, the nucleosome is never forced to leave the DNA completely, but is instead shifted perhaps as much as 100 bases backwards from its original position. How this movement is affected by the histone modifications and additional proteins present *in vivo* will doubtless provide fertile ground for future investigations.

One of the major insights offered by Watson and Crick's discovery was the ba-

sis of the mechanism by which DNA was replicated, but how does the eukaryotic cell, with its thousands of replication sites, ensure that all the information is copied once and once only. For a long time, workers on this topic have debated whether defined origins of replication are involved, as they are in bacteria, but only now count Arturo Falaschi (ICGB, Trieste) reported that in one area at least of the wild-type human genome replication can be shown to begin in a single defined phase.

Phenomena such as exon shuffling imply that genomes are constantly being rearranged, and are not mere static repositories of information. The central contribution of extrachromosomal elements such as plasmids and episomes to this process was elegantly summarized by Stanley Cohen (Stanford). Even in mammalian evolution, transposons and retroviruses have played an important role; Maxine Singer (Carnegie Institution) presented evidence that one of the 4,000 LINE-1 retrotransposons in the human genome still encodes an active reverse transcriptase.

On other occasions, the genome seems to rearrange itself for its own ends. The most prominent example occurs in the immunoglobulin genes, where one of the many available gene segments encoding a variable region must be brought to a position next to the remainder of the gene before it can be expressed. At the κ locus, this can entail either deletion or inversion of a DNA segment, depending on which of the 24 functional variable genes are used, explained Hans Zachau (Institut für Physiologische Chemie, Munich). In all cases, however, the broken ends of the DNA on either side of the initial cut are apparently sealed so as to form hairpins, as Martin Gellert (NIH) showed, before they are nicked to form the final joint (a process reminiscent of the reaction mechanism employed by topoisomerases).

How is the information stored in the DNA used to direct the development of a complete organism? Intensive NMR studies in collaboration with Kurt Wüthrich have led Walter Gehring (Biozentrum, Basel) to conclude that the DNA binding specificity of the homeobox genes he discovered (which are transcription factors specifying segmental identity in metazoans from insects to humans) is determined not only by the 'recognition helix' that sits in the major groove of the DNA, but also by a flexible segment at the amino-terminal end of the homeobox which appears to wrap around the DNA so as to contact the minor groove.

Plants, too, can have homeotic mutations, noted Heinz Saedler (MPI, Cologne), although in the case of mutations affecting flower formation in *Antirrhinum majus*, at least, the proteins in question are related to the mammalian serum response factors, rather than containing homeoboxes. In this case, however, mutations affecting the ex-

pression of any one factor invariably affect two adjacent parts of the flower.

Transformations of another sort are produced by mutations in the retinoic acid receptors. Pierre Chambon (Institut de Chimie Biologique, Strasbourg) described an ambitious set of experiments aimed at elucidating their role by knocking out each of the eight isoforms in separate lines of transgenic mice. To his initial astonishment, lines lacking the $\alpha 1$, $\beta 2$ and $\gamma 2$ receptor all appeared quite normal, and it was only when he had made lines lacking two isoforms at once that clear effects were seen (thank God, he added). The apparent ability of these receptors to substitute for one another is strange in view of their strong evolutionary conservation; perhaps, mused Chambon, their absence results in a phenotype that cannot easily be measured in the laboratory.

Although the discovery of the double helix made it possible to understand how information might be stored in DNA, revolutionary advances in areas such as nucleotide sequencing and computing have been needed before people can contemplate recovering the stored information on a large scale. One of the most exciting parts of the symposium was the account by Daniel Cohen (Centre d'Etudes du Polymorphisme Humain, Paris) and Jean Weissenbach (Institut Pasteur) of their recent progress towards a physical and genetic map of the entire human genome. ("French science in its grandest fashion", remarked Watson later, adding that the audience should not believe what some "worse than third-rate authors" had written elsewhere.)

Already, Cohen has isolated more than a thousand contigs, each containing an average of 15 yeast artificial chromosomes (YACs) and covering more than 70 per cent of the genome, while the number of markers on the map that Weissenbach and his colleagues reported last year has since been more than doubled and is set shortly to double again. Although strategies to overcome problems with chimaeric YACs and marker-poor regions are still being worked out, it is clearly only a matter of time before a combined version of the two maps will be available to guide efforts to sequence the human genome.

Access to this flood of information about ourselves will eventually transform medical practice, and the first signs of this are already apparent. Jean-Louis Mandel (Institut de Chimie Biologique, Strasbourg) described how diseases such as fragile X syndrome and myotonic dystrophy (recently joined by spinobulbar muscular atrophy and Huntington's disease) are caused by a progressive expansion in successive generations of a triplet repeat in the gene. But although this discovery has made accurate diagnosis much easier, the way in which the expansion occurs remains to be understood.

Elsewhere, workers are trying to alter the body's response to disease by adding information to its cells. Michael Blaese (NIH) described studies in which T cells transformed with a retrovirus encoding adenosine deaminase (ADA) were used to correct the crippling immunodeficiency produced by a lack of this enzyme.

Blaese and his colleagues have also shown that a mouse glioma can be completely ablated by treatment with acyclovir following a stereotactic implant of fibroblasts infected with a retrovirus copying the herpes virus gene for thymidine kinase. Similarly, Max Birnstiel (IMP, Vienna) described studies in which a newly developed and very efficient technique for gene transfer was used to induce tumour cells to synthesize high levels of interleukin 2, so as to make them more conspicuous to the immune system. Mice immunized with two doses of such cells were protected against a subsequent lethal challenge with the tumour, and experiments in which the immunization is used to induce a response against an established tumour are under way.

But it was left to Watson to sum it all up. We are now, he said, in the sixth phase of genetics, the phase of the human genome sequence, which had succeeded the ages of mendelian genetics, DNA, the genetic code, the control of genes, and recombinant DNA, respectively. Although this "slightly megalomaniac project" had been started on the strength of charitable donations, future progress should be funded by governments, if only because the project would otherwise be open to exploitation by private companies and their insatiable desire for secrecy.

The sixth phase would, he predicted, give way to a seventh age of genetic prediction. This would make possible the routine diagnosis of vast numbers of genetic conditions, which should be eliminated where necessary by abortion; the world must shed the idea that this is evil, as it is an act of true moral cowardice to allow children to be born with known genetic defects. Although some of these defects might eventually be treatable by gene therapy, this would never be possible for all of them, and prevention is in any case better than cure.

What of germline gene therapy? The consensus is that the world is not yet ready for it, and it may in any case never be necessary. Of course, it is possible to imagine evil-minded dictators abusing it to produce nightmares, but if it could be used to enhance our resistance to disease, or the survival of humanity depended on it, then it would be time to think again. No principle is worth extinction. Watson ended by firmly endorsing the power of genetics to improve the human condition; a lot of people will dislike it, he added, but when he was young a lot of people had disliked Roosevelt — so what! □

Nicholas Short is Biology Editor of *Nature*.

The case for the Higgs boson

Mr William Waldegrave, Britain's cabinet minister in charge of science, has offered a bottle of decent champagne to one who can tell him why the Higgs boson is worth finding (see page 781). These notes may be helpful.

Do not mention gauge invariance, non-Abelian or otherwise. Nor should you begin with a reference to spontaneous symmetry breaking. Of course, there are now plenty of neat proofs that the ground state of a quantum system does not necessarily enjoy the full symmetry of the system as a whole. The obvious case is that of a ferromagnet cooled below the Curie temperature, when the rotational symmetry of the system may be lost by magnetization.

Waldegrave is both inquisitive and intelligent. That he takes his job seriously is borne out by his willingness to worry about the Higgs boson. Evidently the question has arisen in discussions over what should happen next at the European Laboratory for Particle Physics (CERN). With a British director (Dr Christopher Llewellyn-Smith) now at Geneva, Waldegrave cannot just walk away from the plan to install in the LEP tunnel the Large Hadron Collider, part of whose promise is success in the hunt for the Higgs boson. It is uncharacteristically diligent of a minister to seek precisely to understand what money spent will accomplish. But for all those virtues, it would be ridiculous to expect that he would be enlightened by an account of gauge invariance or symmetry-breaking. Avoid both concepts.

Analogy is the only way to start, and the electron and the electromagnetic field the only place. That approach has also the benefit of introducing several British names of which ministers other than Waldegrave will have heard. (Faraday is now on the £20 banknote.) So begin along these lines: *With respect, Minister, it's to do with explaining action at a distance. Newton worried about that when his theory of gravitation required apparently instantaneous interaction between two distant objects. By Faraday's time, the problem was more acute; electrically charged objects were known to repel or attract each other, as did magnetic poles. Faraday began talking of magnetic fields (and you'll remember the school experiments with iron filings around bar magnets). Later, Faraday showed that moving charges made magnetic fields and so we had the electromagnetic field. By the time of Maxwell, it was clear that the electromagnetic field was the means by which electrical charges and magnetic dipoles act on each other 'at a distance'.*

Having begun in this historical vein, you had better stay with it. If you consider that the recitation of the names of British

heroes will advance your cause, you might even interpolate:

To begin with, people thought that electric charges could be arbitrarily big or small, but Faraday made the crucial step of showing that charged atoms, or ions, are always associated with a fixed amount of charge or some multiple thereof, suggesting atoms of electric charge. Then J. J. Thomson showed that these atoms of charge exist; we call them electrons. So far as we know, there are no corresponding atoms of magnetic charge.

It would be unwise to elaborate on the last point. This is hardly the place for a dissertation on magnetic monopoles; indeed, with luck, you can hope to avoid a single mention of the Grand Universal Theories (called GUTs) that predict them. But you cannot avoid explaining how the coming of relativity and then of quantum mechanics influenced thinking on the electromagnetic field. Try something like this: *Relativity fitted naturally with Maxwell's electromagnetism, explaining for example why the electromagnetic field around a charge apparently at rest is purely electrostatic, but that around a moving charge (or an electric current) has magnetic components as well. (Mention of H. A. Lorentz at this point would be courteous, Hollander though he may have been.) But the coming of quantum mechanics was the very devil. For a decade until the late 1930s, people could do no better than to regard the electron as an empirical fact. Only after Dirac had predicted that there must be positrons as well as electrons (verified in 1932) did he and others set out to see whether the charge and the mass of an electron can be derived from first principles. He did not get far.*

Now comes the difficult part. The problem is how to explain the quantization of the electromagnetic field to one who cannot instinctively write down Maxwell's equations (which is no shame), let alone say what they mean. Hand-waving is unavoidable. A few references to the successes of quantum mechanics in the 1930s might be appropriate; in retrospect, Heitler's quantum theory of radiation is especially prescient. But you may be running out of space. So try this:

Dirac was defeated by the complexity of the problem. The quantum theory of an atom requires that each of its components, its electrons for example, should be dealt with separately, but at worst there is only a finite number of them. But the

electromagnetic field has an infinite number of variables, one for each point in spacetime. Finishing Dirac's problem was everybody's goal in the 1940s. Feynman, Schwinger and Tomonaga won the prize. The upshot was a picture in which electrons interact by exchanging photons with each other. But they are not real photons at all, but virtual photons created out of the vacuum.

There's now a dilemma. Do not raise awkward questions such as whether quantum electrodynamics, which calculates an electron mass by subtracting infinities by the technique of renormalization, attains Dirac's goal of deriving the properties of the electron from first principles. Most think it does not. In any case, you've almost used up all your space without mentioning the Higgs boson: best do so quickly: *The Higgs boson, Minister, plays much the same role as the photon, but in a different context. Rutherford's school saw the difference between α and β radioactivity. The second kind, in which nuclear particles give off electrons, forms a bridge between nuclear matter (neutrons for example) and non-nuclear matter such as electrons. In other words, nuclear and non-nuclear matter act on each other. It's action at a distance all over again. So there's another field filling the whole of space that must be quantized. What that exercise predicts is that there should be particles mediating the influence of that field just as the photon mediates the electromagnetic field. They've been found, at CERN. And then there's the last of them, the Higgs boson. That's why its discovery is the last goal for the time being of particle physics.*

That, sadly, is all you have room for on one side of a sheet of A4 paper. Do not be downcast that you have been economical with the truth. Not to have mentioned the strong interaction and the quarks and gluons thereby made necessary is barely forgivable, the omission of neutrinos and the lepton generations other than the electron (the μ and τ particles) is less so.

You may also cringe a little at your conclusion, knowing as you do that the correctness of the electro-weak theory has been essentially verified by the observation of neutral-current processes as well as by the discovery of the $W^\pm$ and Z^0 particles. The real interest of the hunt for the Higgs particle is that it may not be quite what is expected. But best not say that too loudly.

John Maddox

A death in the life of p53

D. P. Lane

THE papers on pages 847 and 849 of this issue^{1,2} establish the involvement of p53 in the induction of apoptosis by radiation and chemotherapeutic DNA-damaging drugs. This discovery brings together two rapidly advancing fields by demonstrating that a tumour-suppressor gene plays an essential role in the induction of programmed cell death.

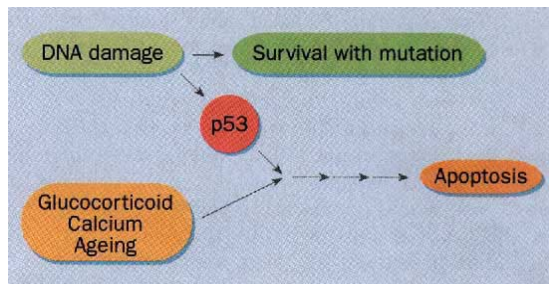
Apoptosis was first defined in the 1970s³. A complete molecular description of this process of programmed cell death initially proved to be elusive, but its central part in normal development has now been clearly established; the phenomenon is recognized to be widespread in multicellular organisms, and underlies such processes as organogenesis, tissue homeostasis and the 'editing' of the immune system to remove autoreactive clones⁴. As well as being induced by developmental control processes, apoptosis can also be induced by toxic insults and, in particular, by agents that damage DNA. It is frequently seen in tumours and may play a key part in the kinetics of tumour growth. Many genotoxic cancer treatments may also exert their effect through the enhanced induction of apoptosis⁵.

Elegant work on the nematode has identified a number of genes responsible for regulating the process⁴, and in mammalian systems two proto-oncogene products have been shown to act as powerful regulators in the pathway. The *myc* gene product, when over-expressed in cells deprived of growth factors, induces apoptosis⁶, while over-expression of the *bcl-2* gene makes cells resistant to it⁷. There was great interest in the observation that over-expression of wild-type p53 could induce apoptosis in a range of cultured cell systems^{8,9}. But the interpretation of these findings was not straightforward, because high levels of p53 induce growth arrest without apoptosis in other cells¹⁰.

The idea that p53 may be an essential component of the apoptotic pathway suffered a set-back when it was found that mice without any p53 gene were viable¹¹. The animals develop normally, and are fertile and immunologically competent, which would seem to be at variance with the idea that p53 is involved in these apoptosis-associated processes. But the demonstration that high levels of p53 are induced by DNA damage has led to the concept that p53 might be part of a damage-control pathway rather than the pathways that are at

work in normal development^{12,13}.

The new studies provide compelling evidence in support of these concepts. Lowe *et al.*¹ and Clarke *et al.*² decided to investigate the role of p53 in a classic system, the induction of apoptosis in thymocytes¹⁴. These immature T cells are very susceptible to apoptosis and are normally short-lived. They can be in-



Two independent pathways lead to apoptosis in thymocytes. One, initiated by DNA damage, absolutely requires the product of the p53 gene; the other, initiated by glucocorticoids, calcium ionophores or ageing, has no such requirement. Inactivation of the p53 pathway will result in the survival of cells exposed to mutagens, and may thus have a pivotal role in the development of cancer.

duced to undergo apoptosis at an even more rapid rate *in vivo* and *in vitro* by exposure to low doses of ionizing radiation or other DNA-damaging treatments, or by treatment with glucocorticoids or by a calcium ionophore in the presence of phorbol ester. Calcium ionophore treatment is thought to mimic the natural signal responsible for the editing out of self-reactive T cells in the thymus.

The two groups have repeated the work of Donehower and his colleagues¹¹ and produced new strains of p53 knock-out mice. Thymocytes from the knock-out mice show a normal apoptotic response to treatment with glucocorticoids but are extraordinarily resistant to the induction of the process by radiation, both *in vitro* and *in vivo*. The p53 null thymocytes are resistant to more than 20 grays, while the control thymocytes show clear induction of apoptosis at doses of only 1 gray. So p53 is essential for the apoptotic response to the radiation-induced signal, but has no part at all in the response to the glucocorticoid-induced signal.

These results are remarkably clean, and underscore the enormous power of the gene knock-out approach, and they show that there are at least two independent pathways to the induction of apoptosis. Moreover, the knock-out mice will allow a crucial examination of the p53

pathway in the induction of apoptosis in other tissues and in response to a wide range of different stimuli (including agents used in cancer chemotherapy). Indeed, Clarke *et al.*² show that p53 may play an important role in cell death resulting from treatment with the topoisomerase 2 inhibitor, etoposide, but that it appears to have no involvement in the spontaneous 'ageing' apoptosis associated with prolonged culture of these thymocytes.

Both groups are able to show a clear effect of gene dosage; that is, cells from heterozygous mice which contain only one copy of the p53 gene are slightly more resistant to radiation-induced apoptosis than are homozygous normal mice with two copies (which are highly sensitive), while the homozygous null mice with no active copies are highly resistant. A gene-dosage effect of this kind has important consequences for the treatment of people carrying germ-line mutations in the p53 gene, such as those affected by the Li-Fraumeni syndrome¹⁵. It also bears upon our understanding of tumour progression, because it implies that cells in which one p53 allele has been somatically mutated may have a survival advantage in the presence of genotoxic agents (for example ultraviolet exposure of skin) over those that have two intact copies of the gene. This in turn could make them more likely to survive a second mutation at the p53 locus (which is frequently seen in cancer cells), or indeed at any other locus.

The mechanism by which p53 makes cells susceptible to radiation-induced apoptosis is not yet clear. Certainly both groups find that p53 levels rise following radiation, and this could specifically activate transcription of apoptosis-inducing genes¹⁶. Alternatively, p53 may interact more directly with the repair machinery, perhaps increasing its fidelity but delaying repair so that lesion-induced

1. Lowe, S., Schmitt, E. M., Smith, S. W., Osborne, B. A. & Jacks, T. *Nature* **362**, 847–849 (1993).
2. Clarke, A. R. *et al.* *Nature* **362**, 849–852 (1993).
3. Kerr, J. F. R., Wyllie, A. H. & Currie, A. H. *Br. J. Cancer* **26**, 239–245 (1972).
4. Raff, M. C. *Nature* **356**, 397–400 (1992).
5. Dive, C., Evans, C. A. & Whetton, A. D. *Semin. Cancer Biol.* **3**, 417–427 (1992).
6. Evan, G. I. *et al.* *Cell* **69**, 119–128 (1992).
7. Vaux, D. L., Cory, S. & Adams, J. M. *Nature* **335**, 440–442 (1988).
8. Yonish, R. E. *et al.* *Nature* **353**, 345–347 (1991).
9. Shaw, P. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **89**, 4495–4499 (1992).
10. Michalovitz, D., Halvey, O. & Oren, M. *Cell* **62**, 671–680 (1990).
11. Donehower, L. A. *et al.* *Nature* **356**, 215–221 (1992).
12. Kastan, M. B., Onyekwere, O., Sidransky, D., Vogelstein, B. & Craig, R. W. *Cancer Res.* **51**, 6304–6311 (1991).
13. Lane, D. P. *Nature* **358**, 15–16 (1992).
14. Wyllie, A. H. *Nature* **343**, 555–559 (1980).
15. Vogelstein, B. *Nature* **348**, 681–682 (1990).
16. Kastan, M. B. *et al.* *Cell* **71**, 587–597 (1992).
17. Branch, P., Aquilina, G., Bignami, M. & Karran, P. *Nature* **362**, 652–654 (1993).

apoptotic signals are sustained. Such an apparently paradoxical phenomenon is seen in cells acquiring resistance to alkylating agents by virtue of defective mismatch repair¹⁷.

The induction of apoptosis by p53 following genotoxic insult may act as a defence mechanism to protect the organism from the propagation of cells that have sustained mutation. The pathway will clearly be modulated by other genes

that regulate apoptosis. Abrogation of this p53 pathway is the most common specific alteration in human cancer, and may be central to the progression of the disease and to its response to treatment by radiation and chemotherapeutic drugs. □

D. P. Lane is in the Cancer Research Campaign Laboratories, University of Dundee, Dundee DD1 4HN, UK.

VOLCANOLOGY

Geochemical hazard indicators

David M. Pyle

THE development of successful models for predicting volcanic eruptions hinges on answering two questions: how long do volcanoes live, and how much magma do they store? On page 831 of this issue¹, Chen and colleagues report the discovery of a correlation between lava chemistry and eruption volume at Unzen volcano, Japan. They find significant differences in the neodymium isotope compositions of small and large eruptive sequences over the past 300,000 years. This suggests that the isotopic signature of lava from early in an eruption is a predictor of the volume still to be erupted.

In this case, with hindsight, this suggestion turns out to be true. Lava has been erupting from Unzen since mid-1991, causing much death and destruction in the process. Samples from early in this eruption have the isotopic signature of previous large eruptions. Had this volume-composition correlation been known in May 1991, geologists could safely have predicted that the eruption would last for at least another few months, and would erupt at least twice as much magma again. This is the first predictive tool to appear that is based on the chemistry of the erupted lavas. The significance of the results extends beyond Unzen, as it reinforces a volume-composition pattern found across the western United States^{2,3}, and suggests that Nd-isotope tracers might be generally applied to continental effusive volcanism in the effort to minimize volcanic hazards.

There is a convincing explanation for this pattern⁴. Magma is believed to be stored in crustal reservoirs before eruption. These reservoirs behave as imperfectly elastic containers, expanding and contracting to accommodate fluxes of melt. They play the role of a chemical filter⁵, modifying the incoming melt by storing and mixing magmas before eruption, and allowing melts to interact with the chamber walls. Fortunately, the incoming melts and the chamber walls

often have significantly different isotopic compositions, so that their respective influences can be detected. Whereas resupply with fresh melt produces erupted liquids with more 'source-like' compositions, prolonged storage and crystallization coupled with wall-rock assimilation generates liquids with more 'crust-like' compositions.

The proposed explanation¹ of the volume-composition relation is that it is controlled by the recharge rate: rapid recharge leaves little time for assimilation, and produces large volumes of mantle-like melts. With lower recharge rates, storage and assimilation become relatively more important, and the resulting eruptions are both smaller and more crust-like. However, the maturity of the plumbing system and reservoir size may also play a role. At some volcanoes², the most crust-like magmas are produced immediately after caldera collapse events, when the chamber roof collapses and melts that have failed to erupt are swamped with crustal contaminants. Later magma, following the same route to the surface, encounters a barren conduit system, already stripped of contaminants. There is also the possibility that as the volume of stored melt increases, and the ratio of volume to surface area in the reservoir increases, so wall-rock contamination becomes less efficient.

The other important questions are how long reservoirs survive, and how long they take to develop. By a quirk of nature, many large-volume rhyolite systems harbour exceptional melt Rb/Sr ratios^{2,3} (in the range 100–1,000) and consequently experience rapid changes in Sr-isotope composition as radioactive ⁸⁷Rb decays to ⁸⁷Sr. With routine measurements of Sr-isotope ratios possible to a precision of a few tens of parts per million, time differences of only a few thousand years between crystallization and eruption ages become detectable. Studies of the 700-km³ Bishop Tuff eruption and its precursors from Long

Valley, California^{3,6} reveal distinct Sr-isotope differences between crystals and their host liquids and between melts erupted at different times.

Modelling⁶ leads to the startling conclusion that melts may have been stored for as long as 300,000–500,000 years before eruption. Although this concurs with suggestions⁷ that magma in caldera-forming systems is produced at rates of around 1 km³ per thousand years, it is not clear that this is a reasonable time-scale^{3,6,8}. Arguments turn on whether the whole chamber had to remain molten, and, if this was the case, on the plausibility of supplying an appropriate heat flux to offset cooling. If Long Valley magma chamber was recharged, then to balance the heat losses of 10⁸ watts requires a magma flux of 1 km³ per hundred years sustained for the lifetime of the chamber⁶. This flux is comparable with the eruption rates of many continuously erupting basaltic and andesitic volcanoes, and is not obviously implausible. Why the chamber should have decided to evacuate after half-a-million years of stability remains a matter of conjecture.

At the other end of the eruption size-scale, time-series analysis of the rates of change of chemical signatures in lavas from continuously erupting volcanoes is now being used to reveal the storage times of magmas before eruption⁵. Results suggest that melt residence times beneath basaltic volcanoes are only of the order of tens of years, and stored melt volumes are consequently small (1 km³ or less). A similar approach has been applied to marine records of explosive eruptions in the Bay of Naples⁹. Eruptive and chemical cycles are correlated, and apparently both have periodicities of around 23,000 years. This timescale is similar to the Earth's orbital precession period, which also appears in climate records, suggesting that external factors may modulate the ascent of magma to the surface.

Although these examples represent important advances in the application of chemical tools to tracing magma systems, the answer to what controls magma-chamber lifetimes is not yet clear.

1. Chen, C. H., DePaolo, D. J., Nakada, S. & Shieh, Y. N. *Nature* **362**, 831–834 (1993).
2. Hildreth, W., Halliday, A. N. & Christiansen, R. L. *J. Petrology* **32**, 63–138 (1991).
3. Halliday, A. N. *et al.* *Earth planet. Sci. Lett.* **94**, 274–290 (1989).
4. DePaolo, D. J., Perry, F. V. & Baldrige, W. S. *Trans. R. Soc. Edin.* **83**, 439–446 (1992).
5. Albarède, F. *Geochim. cosmochim. Acta* **57**, 615–621 (1993).
6. Christiansen, J. N. & DePaolo, D. J. *Contr. Miner. Petrol.* **113**, 100–114 (1993).
7. Smith, R. L. *Geol. Soc. Am. Spec. Pap.* **180**, 5–27 (1979).
8. Sparks, R. S. J., Huppert, H. E. & Wilson, C. J. N. *Earth planet. Sci. Lett.* **99**, 387–389 (1990).
9. Paterne, M. & Guichard, F. *J. geophys. Res.* **98**, 1861–1873 (1993).
10. Sparks, R. S. J., Sigurdsson, H. & Wilson, L. *Nature* **267**, 315–318 (1977).

If some continuously erupting, small-volume systems have short residence times, while other intermittent, large-volume systems are long-lived but continuously fed, this suggests that factors other than magma supply control eruptions. However, the inference that at Unzen recharge rates do control eruption size, and the existence of well-

documented examples where supply of fresh magma demonstrably caused eruption¹⁰ implies that if there is a general pattern to be discovered, it will not be a simple one. □

David M. Pyle is in the Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK.

AGRICULTURE

Better cycling in China

Peter D. Moore

CAN the ecosystem concept be usefully applied in situations where human activity is dominant? It certainly can, as shown by studies of energy flow and nutrient cycling ranging in scale from the mowing of lawns to the energetic balance of the entire Earth. But in the *Journal of Applied Ecology* (30, 86–94; 1993) J. Y. Guo and A. D. Bradshaw now describe an unusually detailed and economically valuable application of the approach to a farming system in China, in which the interactions are complex and the management options are consequently varied.

The study of agricultural systems in ecosystem terms is usually limited in value because of the relative simplicity of the energy and nutrient flow patterns when compared with more natural habitats. This simplicity is, of course, in many respects the precise object of their management, because energy and nutrients are thereby channelled as directly as possible into human consumption with few side-chains and reduction in the losses involved in transferring energy between trophic levels. But such ecosystems are often open, in that they require heavy inputs of nutrients and energy to maintain populations and productivity of a few selected species of plants and animals. Self-sustaining agriculture must move towards greater internal recycling, which means more nutrient and energy flow routes and more complex interactions within the system. In this respect, the ancient farming systems of some areas of China are of especial interest as they have a high degree of self-sustainability and yet are able to support high densities of human population.

The site studied by Guo and Bradshaw is a village in Jiangsu province, having an area of 155 hectares (ha) and a population of 2,353 inhabitants which has increased by about 0.25% a year over the past 40 years. This farming system has proved self-sustainable for the past 2,000 years and has altered little

except that summer (wet season) rice crops were once combined in the same fields with winter (dry season) legumes, whereas the winter crop is now wheat. Artificial fertilizer application, associated with higher yielding wheat, is also now a feature of the system and is required to cope with the growth in population. Of the land area in the

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Gone fishing (for bigger yields) in Jiangsu province.

village, about 45% consists of fields used for cropping, 21% of fishponds, and 8% of water channels, the remaining 26% being occupied by housing, roads and marginal areas. Over 1,500 pigs are also raised annually, of which more than half are exported from the village.

Guo and Bradshaw carried out analyses of the nutrient (N, P and K) and energy flow patterns within the village and have conducted a variety of experiments to see how these patterns can be modified to result in a greater retention of nutrients and a more efficient return of energy to the human population. Several improvements were tested experimentally, such as the harvesting of grasses from the banks of the fish ponds and feeding them to herbivorous fish. In addition, fish ponds were deepened to find out if light penetration was adequate to allow phytoplankton productivity to be increased. In deepening the ponds from the traditional 0.5 m to 2.5 m, phytoplankton production occurred to depths of 1.6 m and annual production figures of 8.4 tonnes ha⁻¹ were achieved,

which is greater than the grain yield of rice for the same period. The combination of these two modifications to the system resulted in increases in fish yields from 2.6 tonnes ha⁻¹ yr⁻¹ to 12.26 tonnes ha⁻¹ yr⁻¹, an increase of 370%. About one-third of the energy intake of the fish was derived from phytoplankton, so both of the management practices contributed to the boost in fish productivity.

The unproductive channels could also be planted with aquatic vegetation and the crop fed to the pigs. A controlled experiment using the normal commercial pig feed and a mixture containing herbage from an alien aquatic plant, *Justicia americana*, common in the area, resulted in no significant change in pig growth rate. But it did result in a 20% saving on commercial pig feed and also served to recover much of the lost nutrients in the ditches and to reduce eutrophication problems downstream. More than 300 kg N, 80 kg P and 560 kg K were recovered per hectare per year by this method and most of this passed through the pigs and was used as fertilizer for the cropped fields.

The enhancement of internal recycling of nutrients via the pigs, as well as use of mud derived from the pond excavations and waste from animal consumption (fish and pig bones) as fertilizer additives, allowed the grain yields to be maintained while at the same time reducing the demand for external fertilizers by about 25%. Against that, the process of recovering and spreading such wastes is labour intensive.

Overall, the modifications to the management of the system resulted in an increase of about 25% of the total agricultural output, with a change in emphasis towards fish production and export (which makes economic sense because of the high value of fish in the marketplace). Use of marginal areas for grass production and channels for *Justicia* culture effectively enlarged the land base for the population, thus reducing population density from 34 to 19 persons per hectare of agricultural land. More important, it left the people better fed. These developments could provide a breathing space in which the pressure on resources is reduced, but in the long term the system can be sustainable only under conditions of stable population. The ecosystem approach can provide a means of quantifying processes and predicting consequences even in such an inscrutable ecosystem as the Chinese village, but in the end there is a limit to the food that it can supply. □

Peter D. Moore is in the Division of Life Sciences, King's College, Campden Hill Road, London W8 7AH, UK.

Limits to improvements?

Richard Thompson

EFFORTS to improve atomic clocks have come up against an unexpected hurdle. The prospect has been that atoms cooled to microkelvin temperatures, with an arrangement called an atomic fountain, would allow orders-of-magnitude improvements in frequency standards. But only last year, it was suggested that collisions between the chilled atoms might change the frequency standard by a small but significant amount. Experiments by Gibble and Chu now confirm this¹, so that continuous recalibration through extrapolation back to collision-free conditions may be necessary.

It is now more than 25 years since the caesium atomic clock was first used to define the second, replacing the earlier definition based on astronomical observations of the mean solar day. Since then the device has remained essentially unchanged in principle. It consists of a beam of caesium atoms whose hyperfine transition frequency, at about 9.2 gigahertz, is probed by a microwave field using the 'Ramsey technique', where the interaction takes place in two spatially separated regions with a field-free region between them. This scheme, for which N. F. Ramsey was awarded the Nobel prize in physics in 1989, enables the maximum resolution to

would then be reduced, allowing greater accuracy, if no systematic effects become significant. One way of doing this would be to use slower atoms, and laser cooling is ideal for this purpose. In this technique atoms have their velocities reduced by running head-on into a laser beam tuned to their transition frequency. Another possibility, and perhaps a more elegant one, is to use a sample of atoms that have previously been cooled and trapped by laser beams, and to project them upwards against gravity. They then interact with microwaves on their way up and again on the way down after gravity has turned them round. Such an atomic fountain, first demonstrated by Kasevich *et al.*³ in 1989, gives a total interaction time of a large fraction of a second, allowing the possibility of a linewidth some two orders of magnitude smaller than in the conventional arrangement.

It also has the advantage of reducing many systematic effects present in atomic clocks. The unexplored hazard was the effect of collisions between the ultracold atoms, in particular the frequency shifts they introduce.

The study of ultracold collisions has become a fertile subject in itself because the physics of such collisions is fundamentally different from the physics of 'conventional' collisions. The reason is that the de Broglie wavelength of a cold atom (typically 0.2 μm) is greater than the scale of the atomic interaction, making the scattering intrinsically quantum-mechanical in nature. It is only in the past few years that a good understanding of these types of collision has emerged, and it is only a year since Tiesinga *et al.*⁴ made a calculation that suggested that, from the available evidence, the collisional frequency shifts might

limit the improvement available by using a fountain set-up. Collisions in general shift frequencies because of the way they perturb atomic structure, and the quantum effects under these conditions make the interactions significantly harder to calculate. The calculations predicted an unexpectedly large shift of the order of 1 millihertz for an atomic density in the fountain of 10^9 atoms per cubic centimetre. This might not sound very much, but it should be remembered that

the current fractional accuracy is between 10^{-13} and 10^{-14} (between 1 and 0.1 mHz) and that the predicted accuracy available is of the order of 10^{-16} (1 μHz). So an experimental measurement of the shift became a very important priority.

Gibble and Chu's experiments pro-

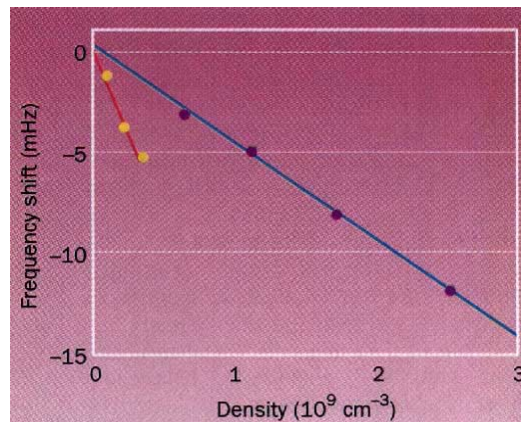


FIG. 2 Frequency shift as a function of atomic density as measured by Gibble and Chu: green line, with all sublevels; red line, with most sublevels removed (see text). Although the atomic beam pulses spread out during a particular measurement, owing to velocity dispersion, calculations show that the mean density quoted here is the important measure.

ceed as follows. First, a sample of 10^9 atoms is loaded into a magneto-optical trap, where they remain for just over 1 second. Then the atoms are projected upwards with a velocity of 2.5 m s^{-1} and cooled further to a temperature of 2.8 μK . They are then optically pumped into the $F=3$ hyperfine level and pass through the microwave interaction region where they experience the first pulse of tunable microwaves centred close to the $F=3$ to $F=4$ transition at about 9.2 GHz. After they have turned round under gravity and interacted a second time with the microwaves, the fraction of atoms in the $F=4$ level is detected and recorded. The spectrum obtained this way (Fig. 1) reveals the linewidth to be just 1.4 Hz, a little over one part in 10^{10} of the transition frequency.

The centre of the resonance can be determined to 1.5 mHz (a thousandth of the linewidth) after just 100 launches. To look for small changes in the transition frequency, it is sufficient to take measurements at two points, one each side of the central peak, roughly half way down the peak. This is what Gibble and Chu did for different densities of the cloud of cold atoms, and the results are depicted in Fig. 2.

They find that the shift is about -4.8 mHz for a density of 10^9 cm^{-3} , with an uncertainty on the density of about a factor of 2; the prediction⁴ for that density was -2.8 mHz . Gibble and Chu are not distressed by the discrepancy,

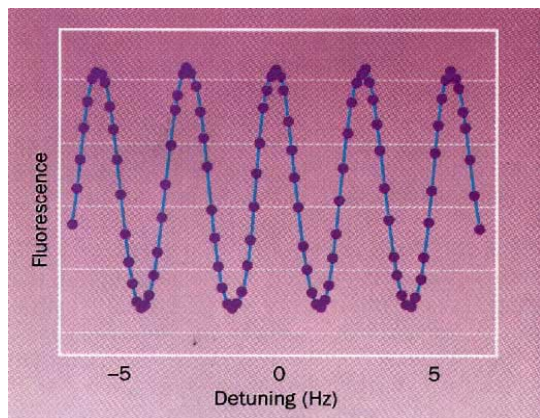


FIG. 1 The central portion of the resonance peak for the $F=3$ to $F=4$ transition studied by Gibble and Chu. The oscillatory fringes are the product of the Ramsey technique which the authors use.

be extracted without the need for a long and highly uniform interaction region. The best caesium clocks achieve a fractional accuracy and stability of between 10^{-13} and 10^{-14} . One of the serious limitations is the total interaction time with the microwaves, limited to a few milliseconds by the speed of the atomic beam and by the separation of the two Ramsey regions².

If the interaction time could be lengthened, the spectral linewidth obtained

because there is fair uncertainty in the interatomic potentials that are used in the calculation.

In fact, the experiment is sensitive to only one sublevel of the F states, although atoms in the other sublevels can contribute to the frequency shifts. As they contribute nothing to the signal, it would seem to be advantageous to remove them from the fountain. The authors achieve this with a small magnetic field (the redundant sublevels are magnetic), but in the event find that the shift is clearly larger, -15.8 mHz at the 10^9 cm $^{-3}$ density already considered (Fig. 2). The theoretical prediction for this case is only -2.2 mHz.

Evidently, ultracold collisions will have to be taken into account if atomic fountains are to be used in frequency standards. A high atomic density is favoured to maximize the signal-to-noise ratio, but if it introduces frequency shifts, the shifts will either have to be well characterized, to allow an appropriate correction to be made, or else an extrapolation to zero atomic density will be necessary. Gibble and Chu point out that for the latter option, it is not necessary to know the absolute density with high precision, but just the ratio of densities. Indeed, they are able to predict the zero-density frequency to a precision of 4×10^{-14} from the present data, and they outline several improvements that can be made to the present apparatus, which already has a stability better than many primary caesium standards. So the ultracold collisions are unlikely to provide a fundamental limit on the fountain experiments, at least in the near future, although they will certainly need to be accounted for in some way.

In the meantime, the study of these collisions is certainly revealing much new and interesting physics. For example, related experiments which measure elastic collision cross-sections of ultracold caesium atoms point to the feasibility of using evaporative cooling of magnetically trapped caesium to reach the temperatures required for the observation of the much-sought Bose-Einstein condensation of atoms⁵. That will open a whole new range of quantum-mechanical effects in cold vapours. □

Richard Thompson is in the Department of Physics, Imperial College of Science, Technology and Medicine, Prince Consort Road, London SW7 2BZ, UK.

1. Gibble, K. & Chu, S. *Phys. Rev. Lett.* **70**, 1771–1774 (1993).

2. Vanier, J. & Audoin, C. *The Quantum Physics of Atomic Frequency Standards* (Adam Hilger, Bristol, 1989).

3. Kasevich, M. A. et al. *Phys. Rev. Lett.* **63**, 612–615 (1989).

4. Tiesinga, E., Verhaar, B. J., Stoof, H. T. C. & Van Bragt, D. *Phys. Rev. A* **45**, R2671–R2673 (1992).

5. Monroe, C. R. et al. *Phys. Rev. Lett.* **70**, 414–417 (1993).

Bacteria forge a new link

Lee Kump

THE list of processes thought to occur abiotically at the Earth's surface has been shrinking on a yearly basis. Increasingly, the involvement of organisms, especially microorganisms, in mediating these processes has become apparent, and Widdel and colleagues (page 834 of this issue¹) now come up with what looks to be a particularly

only minute amounts of free oxygen³. Current models for BIF deposition typically invoke an anoxic deep ocean, enriched in ferrous iron, which provided iron to the surface through upwelling processes⁴. At the surface, the iron was oxidized. It then settled to the seafloor, below the level of wave action, where it accumulated as an enriched iron deposit.

Several pathways of oxidation have been proposed, including reaction in the ocean's surface with low concentrations of free oxygen produced by oxygenic photosynthesizers^{4,5}, and photochemical oxidation of ferrous iron⁶.

Widdel and colleagues propose a third mechanism, the use of ferrous iron as the reductant in light-driven CO₂ fixation by anaerobic bacteria using only photosystem I. Such a pathway is energetically favoured, and its existence is now substantiated by Widdel and co-workers. They obtained enriched cultures of purple non-sulphur bacteria from marine and freshwater muds; and, from stoichiometric arguments involving the amounts of cell mass growth and iron oxidized, they demonstrate that the bacteria do indeed use this pathway for carbon assimilation.

From a strictly biological point of view, this work is notable enough in that it establishes a new category of bacterial metabolism. From a palaeoenvironmental perspective, the finding is highly provocative. Although there have been proponents for a connection between microbial productivity and Precambrian iron deposition^{7,8}, the paucity of microfossils and low organic carbon content of BIFs has been seen by some to constitute strong evidence against biological involvement⁹. Widespread BIFs, such as those of the Hamersley Basin in Australia, would seem to have required massive and persistent microbial blooms to yield the tremendous quantity and uniformity of iron deposition within the basin. Holland⁴ has estimated that the rate of iron deposition in this basin was of the order of 20 mg Fe cm $^{-2}$ yr $^{-1}$. A 4:1 molar ratio of iron oxidized to CO₂ fixed yields an estimate of 1 mg C cm $^{-2}$ yr $^{-1}$ for iron-based primary production in the



Banded iron formation near Ishpeming, Michigan, USA. (Photograph by Preston Cloud, reproduced from *Oasis in Space: Earth History from the Beginning*; Norton, 1988.)

remarkable example. They suggest that one of the most spectacular of supposed abiotic events — the deposition of huge amounts of iron in the seas of Archaean and early Proterozoic times, 3.5–1.8 billion years ago — was in fact biological in origin. The finger of suspicion points to anoxygenic, photosynthetic bacteria, newly discovered by Widdel *et al.*, which use ferrous iron as the electron donor.

That such an organism existed, or exists, has been the subject of much speculation by palaeobiologists and geochemists² interested in the origin of the so-called banded iron formations (BIFs); these are sedimentary ores typically composed of alternating layers of silica- and iron-rich minerals that can often be traced as discrete units over hundreds of kilometres. The reason for such speculation comes from the need for a mechanism of generating BIFs at a time in the history of the Earth when the atmosphere may have contained

Hamersley Sea. Productive regions of the world's oceans today¹⁰ generate about 20–50 mg C cm⁻² yr⁻¹, so the Hamersley Sea need only have been moderately productive to generate the iron deposits preserved there today.

But were such ecosystems able to generate iron deposits on this scale without preserving significant quantities of organic matter in the sediments? The answer has to be 'perhaps'. Sediments underlying moderately productive regions of the open ocean today are typically low in organic carbon and enriched in minerals (in this instance, opal and calcite) produced by pelagic organisms. And in the case of the upside-down Archaeal biosphere proposed by Walker¹¹, a counterintuitive relationship between high productivity in the overlying water column and organic-poor sediments below arises, because the supply of oxidant (ferric iron) to the sediments exceeds that of reductant (organic matter).

Although the observations of Widdel *et al.* are convincing, they are not conclusive. Many questions must be tackled before anaerobic, iron-coupled photosynthetic organisms will be accepted as the generators of the Precambrian BIFs. Were these organisms pelagic or benthic? How did they avoid coating themselves in iron oxides, which would quickly shut off the bacteria from the solar energy they demand? Is it not likely that an efficient, anaerobic, heterotrophic pathway based on iron reduction would have been established? (If so, then considerably higher rates of primary production would be required to sustain the iron deposition rates calculated for BIFs.) What is a reasonable estimate for primary production in a system based on anaerobic photosynthesis and respiration¹²? Whatever the answers, the new discovery will fuel considerable discussion about the nature of the earliest ecosystems on the Earth, and the involvement of iron in microbial ecology today. □

Lee Kump is in the Earth System Science Center and the Department of Geosciences, Pennsylvania State University, University Park, Pennsylvania 16802, USA.

Isotopes and a smoking gun

William M. White

CYCLES characterize geochemical processes operating at the surface of the Earth. The salts dissolved in the oceans do not reside there permanently, rather they cycle through them as they pass from crystalline to sedimentary rock. CO₂ is continually cycled through the biosphere, atmosphere, hydrosphere, and soil and rock reservoirs. Some geochemists have suspected that a much grander geochemical cycle, one encompassing the entire crust and mantle, is also at work in the planet. According to this idea, oceanic crust and sediment are subducted into the deep mantle to rise again as mantle plumes hundreds of millions of years later.

On page 809 of this issue¹, Woodhead and colleagues report new isotope data on volcanic rocks from seamounts south-east of Pitcairn Island that seem dramatically to confirm this theory. Although isotope studies have always been at the heart of the deep-mantle recycling hypothesis, previous analyses have focused on radiogenic isotope variations (those that result from radioactive decay, such as the isotopic composition of Sr and Nd); Woodhead *et al.* combine such analyses with oxygen isotope analyses, and it is variations in the latter that are particularly significant.

Hotspots

Linear chains of oceanic volcanoes, such as the Hawaiian and Society Islands, that occur when lithospheric plates drift over a 'hotspot', are now widely thought to be the surface manifestation of mantle plumes, columns of hot rock that rise buoyantly from the deep mantle. The volcanism results from decompressional melting as the plumes approach the surface. From just how deep these plumes rise is still not resolved. Both laboratory experiments and numerical simulations have demonstrated that plumes can arise only at a boundary between convective regimes, such as that between the Earth's core and mantle at a depth of 2,900 km. In addition, recent calculations of the total plume heat flux show that it matches the expected heat flux from the Earth's core². These and other observations have led to the obvious speculation that plumes arise at the core-mantle boundary. If, however, the mantle itself is divided into two convecting layers, plumes might rise from the internal boundary between them, perhaps the seismic discontinuity at 660-km depth.

Magma erupted at mid-ocean ridges are derived from the upper mantle. Comparison of their chemistry with that

of magmas erupted on oceanic islands demonstrates that plumes differ chemically from the upper mantle. The cause of this difference has been a central question in solid Earth geochemistry for two decades. Early on, it was thought that plumes came from a chemically unprocessed, or primitive, mantle reservoir³. However, plumes are generally enriched in those elements that are enriched in the Earth's crust, and studies of radiogenic isotopes showed that this enrichment must have occurred hundreds, if not thousands, of millions of years ago. From these observations, Hofmann and I proposed⁴ that plumes consist in part of deeply subducted oceanic crust and sediment.

Subduction

Return of oceanic crust to the mantle is implicit in plate tectonic theory, and Armstrong⁵ recognized very early the potential implications for crust and mantle evolution if continent-derived sediments were subducted when plates converge. Chemical and isotope studies of the magmas erupted along island arcs that occur above subduction zones subsequently established that sediment can be subducted to depths of at least 100 km. Indeed, in last week's issue⁶, Plank and Langmuir showed that there are global correlations between the flux of sediment into subduction zones and the composition of magmas erupted from the associated volcanoes. As Plank and Langmuir point out, only a fraction of the sediment flux seems to be consumed in arc magma genesis, implying that the bulk is recycled into the deep mantle.

Nevertheless, unequivocal evidence for deeper subduction of sediment has proved elusive, and other explanations are possible. In a hypothesis that bears some similarity to ours, McKenzie and O'Nions⁷ suggested that the lithosphere underlying continents might detach itself, sink to a convective boundary layer, and ultimately rise as mantle plumes. Melts that had frozen within this subcontinental lithosphere would enrich it in the same elements that are enriched in the crust, so it has not been possible to distinguish between these theories. Also in last week's issue⁸, Class and co-workers interpret radiogenic isotope data from the Kerguelen plume in just this way. These two possibilities are not mutually exclusive. There are three or four geochemical classes of plumes⁹, and a different process may be responsible for each class.

The significance of the oxygen isotope

1. Widdel, F. *et al.* *Nature* **362**, 834–836 (1993).
2. Walker, J. C. G. *Life Sci. Space Res.* **8**, 89–100 (1980).
3. Kasting, J. F. *Science* **259**, 920–926 (1993).
4. Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton University Press, 1984).
5. Towe, K. M. *Precam. Res.* **20**, 161–170 (1983).
6. Cairns-Smith, A. G. *Nature* **276**, 807–808 (1978).
7. Cloud, P. *Econ. Geol.* **68**, 1135–1143 (1973).
8. Walker, J. C. G. *Nature* **309**, 340–342 (1984).
9. Beukes, N. J. & Klein, C. in *The Proterozoic Biosphere* (eds Schopf, J. W. & Klein, C.) 147–151 (Cambridge University Press, 1992).
10. Riley, J. P. & Chester, R. *Introduction to Marine Chemistry* (Academic, London, 1971).
11. Walker, J. C. G. *Nature* **329**, 710–712 (1987).
12. Towe, K. M. *Nature* **348**, 54–56 (1990).

RÉSUMÉ

Twist or bulge?

How does a globular protein adapt to a mutation that forces a surplus residue into a helix? J. S. Kavanaugh *et al.* (*Biochemistry* **32**, 2509–2513; 1993) have looked at the structure at 1.5 Å of the unstable mutant, haemoglobin Catonsville, which has a glutamate inserted into a highly conserved segment of helix (C helix). One of two things could happen: the ends could untwist to admit the intruder or the helix could bulge outwards. The C helix happens to have the 3_{10} conformation (hydrogen-bonding to the third residue along, rather than the fourth, as in the α -helix), and the answer is that the glutamate generates a bulge; moreover the local helix geometry switches from 3_{10} to α . Kavanaugh *et al.* suggest that such changes in helix geometry could constitute a general mechanism of adaptation to insertions or deletions.

Well spotted

In astronomy, the boast is often that biggest is best. But A. R. Patnaik *et al.* from Jodrell Bank are proud to announce the discovery of the smallest gravitational lens yet (*Mon. Not. R. astr. Soc.* **261**, 435–444; 1993). They show that two specks, separated by just 335 milliseconds of arc in radio maps of part of the northern sky, are in fact duplicate images of the same distant radio source. An intervening galaxy, acting as a weaker gravitational lens than any yet known, has bent the source's radiation to give the two spots and — the clincher in lensing studies — an 'Einstein' ring connecting them. More than any other lens system, the pair should help cosmologists to evaluate the Hubble constant (giving the rate at which the Universe is growing) by directly tracing the geometry of space over vast distances. Perversely, doing this will require the biggest and best of radio telescope arrays.

Roaring trade

Put one way, K. McComb *et al.* (*Proc. R. Soc. B* **252**, 59–64; 1993) demonstrate a direct link between a mammalian discriminatory ability and a fitness benefit. Put another, they find that lionesses with young can tell the difference between the roars of their cubs' fathers and those of unfamiliar males. The experiments, carried out in the Serengeti National Park, took the form of playing recordings of known provenance to females with cubs. The lionesses were unperturbed on hearing resident males of the pride, but instantly responded with protective behaviour when played the roars of unfamiliar males. The fitness benefit of the lionesses' reaction is plain — males entering a new social group or area often kill the existing offspring, so to further their own reproductive ambitions.

data reported by Woodhead and colleagues¹ is that they appear to *require* an upper crustal or sedimentary component in the lavas of the Pitcairn seamounts. The two principal isotopes of oxygen, ^{16}O and ^{18}O , are chemically very nearly identical, but modest fractionation between them can occur when oxygen is partitioned between phases in which it is bound in very different ways. This fractionation decreases with the square of the temperature, and should be insignificant at the high temperatures of the mantle. Furthermore, the atomic environment encountered by oxygen in different mantle minerals and melts is similar. Thus significant fractionation of O isotopes occurs only at the surface of the Earth. Pitcairn seamount lavas have O isotope ratios up to a few parts per million higher than the mantle value, which leads to the virtually inescapable conclusion that they contain a component that was once at the surface of the Earth.

The only real alternatives to the Pitcairn plume containing recycled continental or sedimentary material are that the magmas have assimilated sediment in crustal magma chambers or they have exchanged oxygen during subsequent submarine weathering. In this respect, it is somewhat disturbing that the highest, least mantle-like O isotope ratios occur in lavas that apparently experienced extensive fractional crystallization. The extensive crystallization implies shallow magma storage and hence the opportunity to assimilate surrounding rock. The nature of the correlations between O isotope ratios and Sr, Nd and Pb isotope ratios is not consistent with assimilation, and the correlation between O and Nd isotope ratios almost certainly could not result from weathering. Nevertheless, the great potential significance of these results makes one hope that the authors will soon provide the types of data, such as H_2O analyses and O isotope analyses of phenocrysts, that could completely rule out assimilation and weathering.

As Woodhead and colleagues point out, the strong correlations between radiogenic and O isotope ratios imply recent mixing between the sedimentary and mantle components. This suggests that the subducted sediment somehow survived as a chemically and physically distinct region large enough to avoid obliteration by diffusion. Another intriguing result is the sediment/mantle concentration ratios of Sr, Nd and Pb inferred from the curvature of the radiogenic-isotope–O-isotope arrays. The sedimentary component appears to be richer in Sr, slightly richer in Nd, and poorer in Pb than modern clay-rich marine sediments.

Carbonate sediment would provide a

closer, though still imperfect, match, a reminder that deep mantle recycling and mantle plumes may play a role in long-term climate regulation. Mantle plumes add CO_2 to the atmosphere–hydrosphere system, while sediment subduction removes it. If this sediment were carbonate-rich, it might imply a short cycle time because carbonate deep-sea sediment was rare before 600 million years ago. But estimates of the timescale of deep mantle recycling, initially around 2,000 million years, have been growing shorter anyway.

Heating

The initial estimate was based partly on the time required for radioactive heating to cause the subducted oceanic crust and sediment to become buoyant and rise⁴. It now appears that radioactive heating would be less important than conductive heating from the core, allowing a faster cycle time. If the seismically peculiar D' layer at the base of the mantle is the source of plumes, as some geophysicists now suspect (for example, ref. 10), its volume is such that the present plume flux would exhaust it in a few hundred million years, implying that proto-plume material is quickly cycled through the layer. Class *et al.*⁸ have also concluded that the cycle time is short. They argue that isotopic variations in lavas produced over the past 100 million years by the Kerguelen mantle plume require a cycle time of much less than a thousand million years.

Clearly, there are many issues surrounding the origin of mantle plumes and the fate of subducted sediment that remain to be resolved. Although it might be too much to say that Woodhead and colleagues have found the proverbial smoking gun that settles the issue of deep mantle recycling, they do seem at least to have found a gun. Whether it has been recently fired will only be shown with a few additional laboratory tests. There will be some lingering doubts, but the work of Woodhead and others represents some of the strongest evidence yet for this grandest of geochemical cycles. □

William M. White is in the Department of Geological Sciences, Cornell University, Ithaca, New York 14853, USA.

1. Woodhead, J. D., Greenwood, P., Harmon, R. S. & Stoffers, P. *Nature* **362**, 809–813 (1993).
2. Davies, G. F. *J. geophys. Res.* **93**, 10467–10480 (1988).
3. Schilling, J.-G. *Nature* **242**, 565–571 (1973).
4. Hofmann, A. W. & White, W. M. *Earth planet. Sci. Lett.* **57**, 421–436 (1982).
5. Armstrong, R. L. *Rev. Geophys.* **6**, 175–199 (1968).
6. Plank, T. & Langmuir, C. H. *Nature* **362**, 739 (1993).
7. McKenzie, D. P. & O'Nions, R. K. *Nature* **301**, 229–231 (1983).
8. Class, C., Goldstein, S. L., Galer, S. J. G. & Weis, D. *Nature* **362**, 715–721 (1993).
9. White, W. M. *Geology* **13**, 115–118 (1985).
10. Davies, G. F. & Richards, M. A. *J. Geol.* **100**, 151–206 (1993).

Picture an enzyme at work

Guy Riddihough

OIL and water don't mix. So how does pancreatic lipase, an enzyme that knocks around in the essentially aqueous environment of the small intestine, get to grips with fats and oils and break them down into the glycerol and fatty acids that are readily absorbed by the lining of the gut? Several elements of the answer come together in the paper by H. van Tilbeurgh and colleagues on page 814; this is a sequel to the same group's report of last year (*Nature* **359**, 159–162; 1992), discussed in News and Views in the same issue (page 107).

The catalytic activity of pancreatic

lipase, like that of a number of other lipases, is greatly increased when the enzyme comes into contact with a lipid/water interface — this is the phenomenon known as interfacial activation. van Tilbeurgh *et al.* have now crystallized and determined the structure of a human pancreatic lipase–colipase complex in the presence of mixed phospholipid/bile salt micelles. The results tell us a great deal about the mechanistic basis of interfacial activation.

The presence of the mixed micelles in the crystallization solution induces a

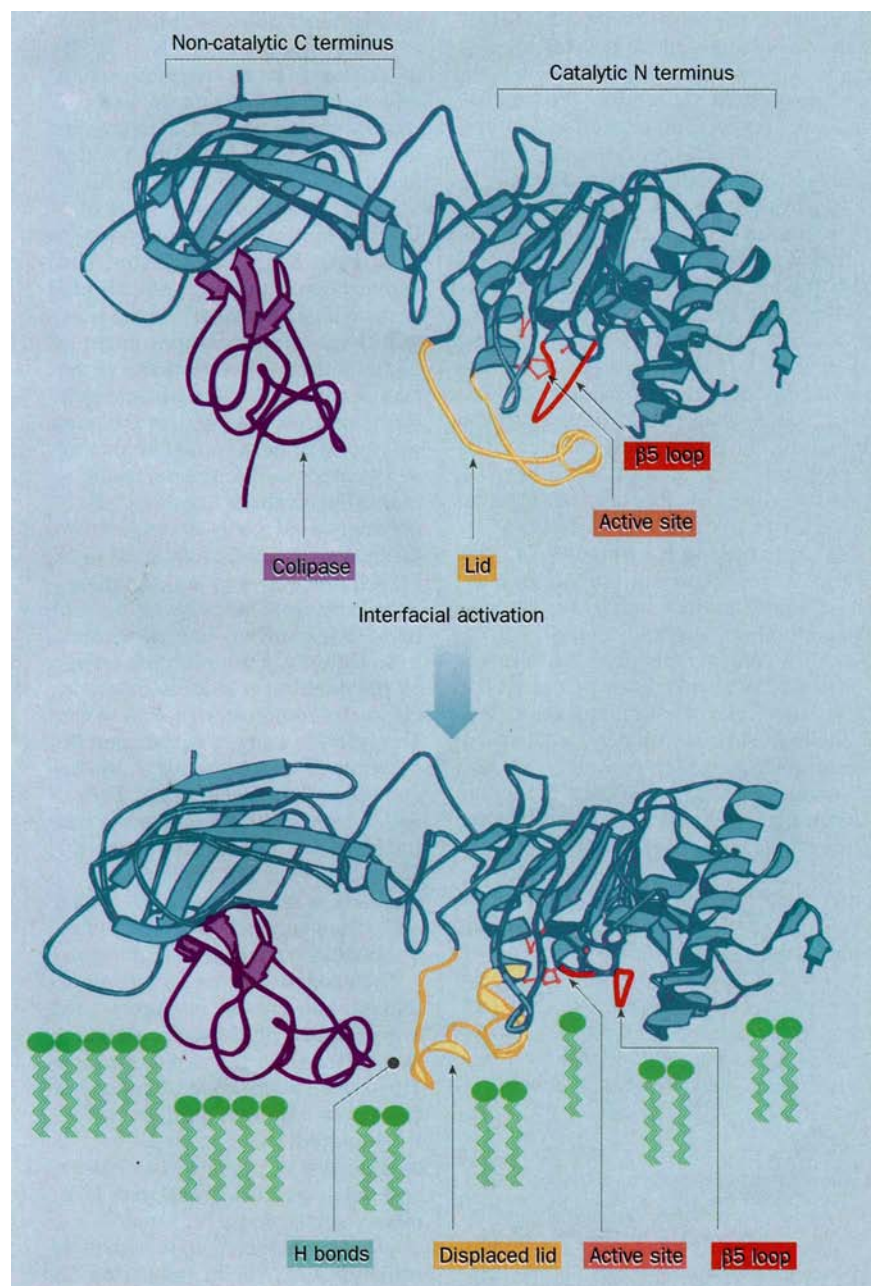
series of conformational rearrangements in the protein complex which, in combination, create a binding site for the lipid substrate. The most dramatic change is seen in the α -helical 'lid' (yellow in the figure), which covers the active site of the enzyme and blocks access to the substrate in the 'closed', inactive, complex. In the presence of micelles (green) the lid pops open. This movement exposes the active site (pink), consisting of a catalytic triad like that of the serine proteinases, which lies at the bottom of a hydrophobic canyon in the catalytic amino-terminal domain of the protein. Furthermore, in the structure there is a region of electron density which could be interpreted as a phospholipid molecule in the canyon, with its two acyl chains sitting comfortably in hydrophobic grooves and the scissile carbonyl carbon positioned in front of the active-site serine γ -oxygen.

A further consequence of the opening of the lid is the restructuring of the β 5-loop (red in the figure). In the closed complex, this loop makes van der Waals contacts exclusively with the lid. When the lid opens these contacts are lost and the loop folds back onto the core of the protein. This creates an electrophilic region, the 'oxanion hole', around the active site serine; the oxanion hole helps stabilize the transition-state intermediate formed during catalysis. The movements of the lid and the β 5-loop expose hydrophobic residues and bury hydrophilic ones on the active-site face of the complex, further increasing the affinity of the complex for the lipid substrate.

The activity of pancreatic lipase is inhibited by a number of components in the small intestine, among them bile salts, secreted by the gall bladder. The inhibitory effect of these components is overcome by colipase, a small protein also secreted by the exocrine pancreas, which binds to the non-catalytic carboxy-terminal domain of the enzyme (purple in the figure).

Colipase has three hydrophobic 'fingers' which point out from the protein complex on the same face as the active site of the lipase. These fingers are thought to enhance binding of the complex to the surface of the micelles, overcoming the inhibitory effect of bile salts. The new structure reveals a second important role for colipase. Once bound, interfacial activation brings the colipase molecule close to the lid and three hydrogen bonds between the fingers and the lid stabilize the conformation of the open complex. This interaction seems to be crucial for the activity of the enzyme. □

Guy Riddihough is an assistant editor of *Nature*.



Acid rain not only to blame

Donald Charles

THE vision of pre-industrial society as ecologically neutral takes a knock on the head on page 824 of this issue¹, with evidence from Renberg and colleagues that prehistoric agricultural practices noticeably altered the pH of lakes in Sweden. The changes were the opposite of those that are causing concern today — acidification — and were smaller. But they demonstrate the role that land use can have in controlling lake chemistry.

As concern has grown over lake acidification in diverse parts of Europe and North America during the past 20 years, the influence of changes in land use has become a prominent part of the debate. How important is land use compared with deposition of strong acids derived from the combustion of fossil fuels? Several hypotheses have it that land-use change, primarily growth of forest or development of heathlands following fires, logging and abandonment of agriculture, could have caused the recent acidification of low-alkalinity surface waters^{2,3}. To an extent, this is backed up by the work of Renberg *et al.*, who provide evidence here and elsewhere⁴ of significant long-term regional effects on surface-water pH, including an increase in pH (increased alkalinity) following conversion of forest to agricultural land and a decline in pH (acidification) following abandonment of agriculture. The full extent of the modern acidification, however, cannot be attributed to changed land use.

Renberg, Korsman and Birks use sediment composition to determine the effects of land-use change wrought by Iron Age peoples in 14 widely separated lakes in southern Sweden (where acid deposition is a severe problem today and a large proportion of lakes are acidified). Sediment cores were dated using ²¹⁰Pb and ¹⁴C; vegetation and land use were determined from pollen; and lakewater pH was inferred from remains of diatoms⁵. The authors' ability to infer water chemistry quantitatively from sediment diatom assemblages is due to rapid developments over the past 5–10 years. It is now possible to infer pH with a standard error of about 0.3 pH units. This new tool has spurred many studies of acidification, and is being applied to other issues such as runaway algal growth and climate change.

Renberg *et al.* reconstructed the lakes' history over their entire lifetime (more than 10,000 years), and they compare the relative importance of the factors potentially responsible for acidification of surface waters. These factors include natural processes such as progressive

leaching from the soil of base cations; removal of forests and replacement by agricultural activities during the Iron Age; regrowth of forests following abandonment of agricultural practices many centuries later; and acid deposition, beginning in the nineteenth century. They find that pH declined from about 7.0 to 5.5 over the first few thousand years following deglaciation, owing to natural processes. After deforestation and expansion of an agrarian economy during the Iron Age, pH increased by 0.5–1.4 units in 12 of the 14 lakes. During the nineteenth and twentieth centuries, pH declined, in the past few decades to less than 5.0 in some lakes, a level toxic to many aquatic organisms. Though it is not possible to determine the relative roles of revegetation and acidic precipitation from this work alone, in conjunction with other studies⁶ Renberg *et al.* argue that the pH decrease to below 5.5 is due to acid precipitation.

In this study, Renberg *et al.* effectively perform natural experiments on lake acidification, with the advantage that the results are 'controlled' in that the same lakes are being examined as are recording the modern industrial assault. The results also cover much longer periods than are possible with modern experimental field studies and they are based on direct evidence of intercomparable past conditions.

The role of land use in modifying lake acidity is important not only in an overall scientific understanding of the process, but more urgently because of the need to control sulphur and nitrogen emissions, primarily from power plants. It is also relevant to management of individual lake-catchment systems. If land-use changes on a regional scale play a larger role in surface-water acidification than allowed for in government assessments, less need would be perceived to control emissions. On the other hand, if the role of land use is overestimated, then emission-reduction policies may be insufficient. □

Donald Charles is in the Patrick Center for Environmental Research, Academy of Natural Sciences, 1900 Benjamin Franklin Parkway, Philadelphia, Pennsylvania 19103, USA.

1. Renberg, I., Korsman, T. & Birks, H. J. B. *Nature* **362**, 824–827 (1993).
2. Rosenqvist, I. T. *Sci. Total Environ.* **10**, 39–49 (1978).
3. Krug, E. C. & Frink, C. R. *Science* **221**, 520–525 (1983).
4. Renberg, I. *Phil. Trans. R. Soc.* **B327**, 357–361 (1990).
5. Birks, H. J. B. *et al. Phil. Trans. R. Soc.* **B327**, 263–278 (1990).
6. Battarbee, R. W. *et al. Nature* **314**, 350–352 (1985).

King Coal, part II

COAL is an uncompetitive fuel. Its high ratio of carbon to hydrogen gives less heat and more carbon dioxide than oil and gas; it burns well only in large elaborate furnaces; it produces ash and sulphur dioxide. Worst of all, it has to be dug out of the ground by expensive skilled labour.

As a chemical feedstock, however, coal is still attractive. Its high proportion of carbon is an advantage here; so is its complex, benzenoid, easily disrupted molecular structure. It used to be the basis of the whole heavy organic chemical industry, and might be again. Daedalus now proposes, not the underground gasification of coal, but its underground polymerization.

Coal is approximately C₁₀H₇O. It must have reached thermodynamic equilibrium over the years, and represents the most stable state for this composition. But a bit of added hydrogen or oxygen could tip the balance. The most stable state of C₁₀H₁₀ could well be polystyrene; that of C₁₀H₈O₄ might be polyester. The copious benzene rings and potential carboxylic acid groups in coal suggest that these transformations might not be too difficult. The methane of the coal seam could also be oxidized to formaldehyde, the basis of the acetal polymers, or even to higher alcohols — the other half of the polyester molecule. All these thermoplastic polymers melt easily. Once formed in place, they could be extracted by the Frasch process — pumping superheated steam down a borehole to blow them to the surface in molten form.

So DREADCO chemists are using all the cunning of modern catalytic chemistry to polymerize coal *in situ*. They plan to pump a suspension of catalysts in steam and air or hydrogen, down into a coal seam. They will launch intense microwaves down the borehole to start the reaction, and will then pump down more reagents to keep it going. The seam will be converted into a vast bed of crude acetal/styrene/ester copolymer.

Thus coal will come into its own. Oil and gas will provide our energy: they are good at that. Coal will provide the raw material for the age of plastics. The crude geopolymer will of course have to be cleaned, purified and fractionated; in the process, its sulphur and nitrogen will be extracted and put to good chemical use. The miner, whose dangerous and unpleasant labour is so misguidedly romanticized, will be eliminated. But coal will still be king.

David Jones

Universal tree of life

SIR — Phylogenetic comparisons of complete small-subunit ribosomal RNA (rRNA) sequences have changed well-established ideas about early events in the evolution of eukaryotes. Nonetheless, the incongruence of rRNA-based phylogenies with molecular trees derived from elongation factors or DNA-dependent RNA polymerases presents a challenge to molecular evolutionists. In rRNA phylogenies the earliest branches are nonphotosynthetic, amitochondriate taxa. They are separated from the more recently diverged kingdoms of plants, animals and fungi by a series of independent protist branches, including *Entamoeba*¹.

Hasegawa and Hashimoto in Scientific Correspondence² suggest that unusual G+C compositions erroneously place the diplomonad *Giardia lamblia* (74.7% G+C) and the microsporidian *Vairimorpha necatrix* (37.5% G+C) rather than *Entamoeba histolytica* (the earliest divergence in elongation factor and DNA-dependent RNA polymerase phylogenies) at the base of the eukaryotic tree. We share their concern about potential misleading effects of biased nucleotide compositions in rRNAs used to infer evolutionary histories. Because there is no accepted theoretical method for compensating for effects of biased nucleotide compositions in phylogenetic inferences, we have sequenced the small-subunit rDNA of the diplomonad *Hexamita inflata*³ (50% G+C) and the microsporidian *Spraguea lophii* (49% G+C). Using distance, parsimony and maximum-likelihood methods, the overall picture of eukaryote small-subunit rRNA phylogeny remains unchanged. Diplomonads, trichomonads and microsporidians represent the earliest diverging lineages, but their relative branching order is influenced by the G+C compositions of prokaryote outgroups. By contrast, *Entamoeba* consistently diverge higher in the tree after the separation of euglenoids/kinetoplasts, acellular slime moulds and amoebflagellates.

Reconstruction of phylogenetic history from molecular sequence data is a probability exercise based on a specific model of genetic change. When various models or different genes are used, statistical measurements can provide strong support for contradictory phylogenies. Deciding between discordant branching patterns frequently reduces to arguments about the "correct model" or "best molecular document" for inferring evolutionary history. In the final analysis, conflicting molecular data sets can be judged by considering the biology of the considered organisms. The fit between trees derived from the small-subunit

rDNA data and morphological and ultrastructure data is unmatched by any other gene used to infer phylogenetic frameworks. When measured by these criteria, the reliability of rRNA-based phylogenies is remarkable and unparalleled in phylogenetic reconstructions of the universal tree of life.

Mitchell L. Sogin

Gregory Hinkle

Detlef D. Leipe

*Marine Biological Laboratory,
Woods Hole, Massachusetts 02543, USA*

1. Sogin, M. L. *Curr. Opin. Genet. Dev.* **1**, 457–463 (1992).
2. Hasegawa, M. & Hashimoto, T. *Nature* **361**, 23 (1993).
3. Leipe, D. D., Gunderson, J. H., Nerad, T. A. & Sogin, M. L. *Molec. biochem. Parasit.* (in the press).

SIR — Hasegawa and Hashimoto correctly point out¹ that rRNA trees might be misleading in defining the evolutionary relationships between very distantly related organisms. They suggest instead that protein trees should be used for this purpose. Indeed, protein trees encompassing the three domains of life (archaeobacteria, eubacteria and eukaryotes) are becoming more popular as the number of available protein sequences from archaeobacteria increases. In particular, several authors have rooted the universal tree of life in the eubacterial branch, using two composite trees of duplicated protein families (elongation factors and membrane ATPases α and β subunits)^{2,3}.

Nevertheless, protein trees could be as misleading as rRNA trees for very distantly related organisms. Well-known mistakes in extrapolating species trees from protein trees have arisen from lateral gene transfers, unrecognized paralogy (duplication of genes before separation of the lineages under investigation) and unequal rates of evolution. We recently summarized⁴ all the data obtained so far by comparing archaeobacterial housekeeping proteins with their homologues from eubacteria and eukaryotes at the sequence level.

As we expected, our analysis revealed contradictions between protein trees and the rRNA tree, and between protein trees themselves. We also find that the composite ATPase and elongation factor trees used to root the tree of life are misleading. First, the finding of V-type ATPases in two eubacteria suggested that V- and F0/F1-type ATPases are paralogous. This was recently confirmed by the discovery of an F0/F1-type ATPase in an archaeobacterium which already harbours a V-type enzyme⁵. As a consequence, the eubacterial rooting previously obtained from ATPase evolution is inconsistent as it was based on

phylogenetic trees in which these two paralogous families were mixed in single trees. Second, cladistic analysis suggests that the elongation factors EF1 α (Tu) and EF2(G) are too divergent to root with confidence the composite tree of their two families, again invalidating the inferred eubacterial rooting.

The difficulty of rooting the universal tree of life using protein trees is also emphasized by our recent analysis of glutamate dehydrogenase phylogeny⁶. Trying to root one subfamily of glutamate dehydrogenase harbouring representatives of the three domains of life, using the paralogous subfamily as an outgroup, we obtained different roots according to the method of tree construction used. Interestingly, the root was located either in the eukaryotic branch or in the archaeobacterial domain, but never in the eubacterial branch.

Caution is therefore necessary in drawing definite conclusions from either rRNA or protein-tree analyses. In particular, it is by no means clear that the problem of rooting the tree of life is now solved.

Patrick Forterre

Nadia Benachenhou-Lafha

Bernard Labedan

*Institut de Génétique et Microbiologie,
Université Paris-Sud, Centre d'Orsay,
91405 Orsay Cedex, France*

1. Hasegawa, M. & Hashimoto, T. *Nature* **361**, 23 (1993).
2. Gogarten, J. P. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 6661–6665 (1989).
3. Iwabe, N. et al. *Proc. natn. Acad. Sci. U.S.A.* **86**, 9355–9359 (1989).
4. Forterre, P. et al. *Biosystems* **28**, 15–32 (1993).
5. Sumi, M. et al. *FEBS Lett.* **3**, 207–210 (1992).
6. Benachenhou-Lafha, N., Forterre, P. & Labedan, B. *J. molec. Evol.* **36**, 335–346 (1993).

Phytoplankton productivity?

SIR — Falkowski and Wilson¹ elegantly demonstrate that historical Secchi disk measurements show no evidence of a significant increase in North Pacific phytoplankton biomass in this century. Unfortunately, they also conclude that increased absorption of anthropogenic CO₂ by phytoplankton in the North Pacific is therefore equally unlikely. This conclusion assumes that phytoplankton productivity can be usefully indexed by biomass levels, which is inappropriate when applied to the North Pacific.

Secchi disk measurements in the North Pacific, while clearly a surprisingly sensitive measure of phytoplankton biomass, cannot be used to infer changes in phytoplankton productivity. Unlike the North Atlantic, phytoplankton biomass in the North Pacific is controlled by zooplankton grazing throughout the year², including the spring bloom. In-

creased phytoplankton productivity may therefore not be reflected in increased phytoplankton biomass levels, but may instead be rapidly transferred to higher trophic levels. Changes in phytoplankton community structure could also result in higher rates of production per unit biomass. Such changes could happen if, for example, the mean generation time of phytoplankton decreased as different groups of phytoplankton become more abundant; population growth rates, and therefore production, are directly proportional to generation time³.

Several lines of evidence point to a change in North Pacific productivity, particularly at higher trophic levels. As well as the doubling of chlorophyll in the central North Pacific gyre between 1965 and 1985 (ref. 4) mentioned by Falkowski and Wilson, zooplankton biomass in the Gulf of Alaska also doubled between the 1950s and 1980s (ref. 5). A similar trend in zooplankton abundance is apparent in the western North Pacific⁶ — although perhaps not in the central North Pacific⁷. The productivity of many fish populations in the northeast Pacific has also shown sharp increases, beginning in the late 1970s (ref. 8). Similar changes in a wide range of physical oceanographic variables have also occurred in the eastern Pacific⁹. These changes may be associated with the intensification of the Aleutian Low¹⁰, a major meteorological system which dominates the weather of the North Pacific region.

The detailed reasons for these increases in productivity at higher trophic levels are unknown. They could involve increased levels of wind-driven mixing of sub-surface nutrients through the thermocline, shifts in phytoplankton community structure towards more intrinsically productive species, or increased insolation.

However, the assumption that productivity must be directly related to biomass or chlorophyll is a fallacy. Particularly in tightly coupled systems such as the North Pacific, increased phytoplankton production may not be reflected in increased phytoplankton biomass levels; production may instead be rapidly exported to higher trophic levels not readily indexed by Secchi disk measurements. There have been striking changes in biological productivity at higher trophic levels in the North Pacific, the mechanisms for which have yet to be fully explained. Nevertheless, the existence of these changes strongly suggests that Falkowski and Wilson's inference that open ocean phytoplankton productivity has remained roughly constant since the industrial revolution is premature, and may even be false. The potential for the oceans to act as an enhanced sink for anthropogenic CO₂ remains,

although the absolute increase in the rate of oceanic CO₂ sequestering possible may still be small relative to the global carbon budget¹¹.

David W. Welch

*Pacific Biological Station,
Department of Fisheries and Oceans,
Nanaimo, British Columbia,
Canada V9R 5K6*

1. Falkowski, P. G. & Wilson, C. *Nature* **358**, 741–743 (1992).
2. Parsons, T. R. & Lalii, C. M. A. *Rev. Oceanogr. Mar. Biol.* **26**, 317–359 (1988).
3. Charlesworth, B. *Evolution in Age-structured Populations* (Cambridge Univ. Press, 1980).
4. Venrick, E. L., McGowan, J. A., Cayan, D. R. & Hayward, T. L. *Science* **238**, 70–72 (1987).
5. Brodeur, R. D. & Ware, D. M. *Fish. Oceanogr.* **1**, 32–38 (1992).
6. Odate, K. *Bull. Tohoku Reg. Fish. Res. Lab.* **48**, 31–47 (1986).
7. Kawamura, A. *Bull. Plankton Soc., Japan* **35**, 175–177 (1988).
8. Beamish, R. J. *Can. J. Fish. Aquat. Sci.* (in the press).
9. Kerr, R. A. *Science* **255**, 1508–1510 (1992).
10. Trenberth, K. E. *Bull. Am. Met. Soc.* **71**, 988 (1990).
11. Sarmiento, J. L. & Siegenthaler, U. in *Primary Productivity and Biogeochemical Cycles in the Sea* (eds Falkowski, P. G. & Woodhead, A. D.) 317–332 (Plenum, New York, 1992).

FALKOWSKI AND WILSON REPLY — Welch objects to our application¹ of a simple algorithm² relating phytoplankton chlorophyll to productivity, calling it a “fallacy” in the North Pacific because of rapid export of production to higher trophic levels. In doing so he is confusing carbon fixation with the fate of phytoplankton³. Obviously chlorophyll is a pool and primary production is a flux, but primary production (carbon fixation) can be quantitatively related to chlorophyll through light, nutrients and temperature^{4–6}, independent of the fate of the production.

We agree that the specific algorithm we used would have been inappropriate if we were interested in examining seasonal or short-term changes in primary production, not because the algorithm does not include a grazing term but because it does not include terms for irradiance and quantum efficiency. But unless there have been marked changes over the past 30 years in insolation, the quantum efficiency of phytoplankton photosynthesis, or rates of nutrient supply, interannual changes or lack thereof, in chlorophyll (which we inferred from interannual changes in Secchi depths) will reflect interannual variations in primary production in the North Pacific or any other ocean basin.

Welch implies that to account for the apparent increases in zooplankton biomass over the past 30 years in the North Pacific without a marked concomitant increase in phytoplankton biomass (as reflected by chlorophyll), phytoplankton productivity must have increased. This would require that for the same amount of light energy (measured incident solar radiation has not varied by more than 0.5% since the beginning of

the twentieth century⁷) more carbon is fixed per unit chlorophyll (the quantum yield has changed). Although the quantum yield of photosynthesis can vary in the ocean⁸, a systematic, multi-decadal change in the quantum yield of twofold is highly implausible. Moreover, if such an increase in chlorophyll-specific production had occurred and was responsible for the changes in zooplankton biomass, the hypothesized increase in primary productivity would have to reflect an increase in new production⁹ which, in turn, must be fuelled by a systematic increase in nutrient supply to the region. If, for example, primary productivity in the region is limited by iron¹⁰, and iron is becoming increasingly available, the iron stimulation would have to lead to a reduction in upper ocean nitrate and phosphate in the summer over the past 30 years. There is no evidence of a systematic decrease in nutrients in the region¹¹. Examination of the data showing the interannual variability in zooplankton biomass at Ocean Weather Station P (ref. 12) suggests there are decadal cycles, but the long-term, systematic changes are not clear.

Finally, as we discussed previously¹, even if primary productivity were to have increased, unless the increase was supported by nutrients external to the ocean, or led to depletion of the excess nutrients in the upper ocean, the effect on the drawdown of anthropogenic CO₂ would be insignificant¹³. We stand by our conclusion that the effect of open ocean primary production on the drawdown of anthropogenic CO₂ appears to be extremely small in comparison with other sinks.

Paul G. Falkowski

Cara Wilson*

*Oceanographic and Atmospheric Sciences
Division,
Brookhaven National Laboratory,
Upton, New York 11973, USA*

* Present address: College of Oceanic and Atmospheric Sciences, Oregon State University, Corvallis, Oregon 97331, USA.

1. Falkowski, P. G. & Wilson, C. *Nature* **358**, 741–743 (1992).
2. Eppley, R. W., Stewart, E., Abbott, M. R. & Heyman, U. *J. Plankton Res.* **7**, 57–70 (1985).
3. Steele, J. H. & Henderson, E. W. *J. Plankton Res.* **14**, 157–172 (1992).
4. Riley, G. A. *J. mar. Res.* **6**, 54–73 (1946).
5. Falkowski, P. G. & Woodhead, A. D. (eds) *Primary Productivity and Biogeochemical Cycles in the Sea* (Plenum, New York, 1992).
6. Kirk, J. T. O. *Light and Photosynthesis in Aquatic Systems* (Cambridge Univ. Press, 1983).
7. Crommelynck, D. *Adv. Space Res.* **9**, 51–61 (1989).
8. Falkowski, P. G., Ziemann, D., Kolber, Z. & Bienfang, P. K. *Nature* **352**, 55–58 (1991).
9. Iverson, R. L. *Limnol. Oceanogr.* **35**, 1593–1604 (1990).
10. Martin, J. H., Gordon, R. M., Fitzwater, S. E. & Broenkow, W. W. *Deep-Sea Res.* **36**, 649–680 (1989).
11. Parslow, J. S. thesis (Univ. British Columbia, 1981).
12. Brodeur, R. D. & Ware, D. M. *Fish. Oceanogr.* **1**, 32–38 (1992).
13. Sarmiento, J. & Orr, J. *Limnol. Oceanogr.* **36**, 1928–1950 (1991).

Soviet meltdown

Zhores A. Medvedev

No Breathing Room: The Aftermath of Chernobyl. By Grigori Medvedev. *Basic Books*: 1993. Pp. 213. \$20, £16.50. UK distribution, HarperCollins.

Ablaze: The Story of Chernobyl. By Piers Paul Read. *Random House/Secker and Warburg*. Pp. 416. \$25, £18.50. UK publication date, 26 May.

THE Chernobyl disaster seven years ago this week (26 April 1986) is still treated as a growing menace in the Ukraine and Belarus. The term "radiation genocide" is commonly used in Kiev and Minsk newspapers to refer to the accident.

Before the collapse of the Soviet Union, the central Soviet budget met the cost of the decontamination, medical treatment and checkups of the several hundred thousand victims and liquidators, as well as the compensation payments to those still living in contaminated areas. But with the Soviet government now gone, both the Ukraine and Belarus have been forced to create special 'Ministries of Chernobyl' and have introduced a 12 per cent 'Chernobyl Tax'.

Last Monday was not only the seventh anniversary of the accident. It was also the closing date for entries to the international competition to design a new Chernobyl 'sarcophagus'. The previous one, hastily built to entomb the destroyed reactor with about 95 per cent of its original nuclear fuel, is now unstable. The new project will convert the site into a long-term ecologically safe system and dismantle, process, transport and bury the fuel and radioactive materials contained in the old sarcophagus. The Ukraine, which announced the competition, apparently expects the international community to foot the bill for the construction of the winning project (at least 30 large companies from the United States, France, Canada, Italy and Germany have submitted entries).

It has become traditional in the West (as well as in the former Soviet Union) to move publication dates of new books on Chernobyl as close to the anniversary of the accident as possible. This year is no exception. But these two books do not contain new scientific or technical information or add anything to existing knowledge of the environmental and medical consequences of Chernobyl. They mostly supplement the picture of the Soviet political system, how it oper-

ated in different circumstances and how it created and controlled its nuclear-energy network.

Grigori Medvedev, a nuclear engineer and the former deputy head of the department of nuclear energy in the



Life goes on, despite nuclear accidents and the collapse of the Soviet empire. Taken from *Chernobyl: Insight from the Inside* by V. M. Chernousenko (Springer, 1991).

Ministry of Energy of the USSR, wrote a book about Chernobyl that was published in Russia in 1989. One of the best on the subject in both substance and style, it proved to be extremely influential, appearing in English as *The Truth About Chernobyl* (Basic Books, 1991). His latest book, written in 1990, is not about the aftermath of the disaster as its English subtitle suggests. It is in fact about the problems Medvedev encountered in trying to publish his original book and some earlier stories about nuclear energy — that is, about his struggle against the complex system of Soviet censorship.

The Truth About Chernobyl was written in 1987 and submitted for publication to *Novy Mir*, one of the best Russian literary magazines. Although the editors realized the importance of the work,

they were unable to publish anything without higher approval. This meant that the book not only had to pass general political censorship, commonly known as *glavlit*, the main obstacle for works of fiction; it also had to be approved by the relevant government institutions, the Ministry of Atomic Energy and the USSR State Committee on the Utilization of Atomic Energy, whose senior officials were named by Medvedev as being responsible for the accident.

The descriptions of the author's visits from one bureaucrat after another take up the bulk of *No Breathing Room*. The delay in publication of the original book would probably have been much longer

had Medvedev not given the manuscript to Andrei Sakharov to read in October 1988. Sakharov greatly admired the work and sent a personal letter of endorsement to Mikhail Gorbachev, then the omnipotent Secretary General of the Communist Party and Chairman of the Supreme Soviet. But Gorbachev was also unable to arrange publication and so he put the problem before the Politburo. It was not solved there either: the Politburo wanted the opinion of the department of nuclear energy of the Central Committee of the Communist Party. The situation seemed hopeless, particularly as the head of this department, Vladimir Marin, was portrayed unfavourably in the book. Only in June 1989, when the highest authority of the land became the newly elected Congress of People's Deputies of the USSR and when the draft of the new law on the press provided for the abolition of censorship, did

Novy Mir publish Medvedev's story for the benefit of its 1,600,000 subscribers.

But it was not an entirely happy ending. In a comprehensive introduction to *No Breathing Room*, David R. Marples not only gives some hard facts about the Chernobyl disaster; he also indicates that the impoverishment of most of the population as a result of soaring inflation caused by 'shock therapy' economic policies, and the removal of subsidies to the publishing trade, made it difficult for high-quality literary magazines to cope, despite the so-called freedom of the press. Many degenerated into sensational tabloids or nationalistic journals. *Novy Mir* today barely survives with only 100,000 subscribers.

Medvedev argues that the Chernobyl disaster and the resulting struggle for truth were the main causes of the abol-

ition of censorship in the Soviet Union. *Ablaze*, a brilliantly written story about Chernobyl that approaches the topic from many different angles, takes these arguments further. The author tries to prove that the disaster was also the main cause of the collapse of communism and the Soviet Union.

Piers Paul Read begins with a short but comprehensive overview of the Soviet nuclear energy programme since the Second World War. He explains how the Soviet research and development system worked at all levels, top to bottom, from Stalin to the millions of Gulag prisoners who built all 15 of the secret Soviet atomic towns that existed in 1953. Subsequent leaders were also devoted to giving nuclear scientists and designers everything they wanted. President Gorbachev wanted to double Soviet nuclear-generated electricity within 5 years — 50 per cent of this rapid growth was to be achieved by the construction of the Chernobyl-type graphite-moderated reactors RBMK-1000 and RBMK-1500. But this programme was ruined by the Chernobyl disaster; and without it *perestroika* had no chance of success. Powerful environmental mass movements emerged in the Ukraine, Belarus and the Baltic states; and they, in turn, were gradually transformed into nationalistic 'people's fronts', whose members won seats in local parliaments and advocated full independence from Moscow and the end of the communist empire.

This link between Chernobyl and the disintegration of the Soviet Union is obvious, but it was, of course, not the only factor in the fall of the superpower. Newly independent states now enjoy sovereignty, but they still need electricity. Sixty per cent of electricity in Lithuania is generated by two RBMK-1500 reactors; 40 per cent in the Ukraine is produced by nuclear power stations; and 50 per cent in Kiev is still supplied by the part of Chernobyl that remains in operation. Armenia is begging Russia to recommission two old reactors that had been shut down for good in 1988 under popular pressure. But after burning all the trees in Yerevan — even those in the botanical garden — as firewood, nuclear-generated electricity seems a better option than the cold. Russia is now the only country among the 15 post-Soviet republics that can design, construct and operate nuclear power stations. The recent decision of the Russian government to revive its nuclear-energy programme indicates that the Soviet common joint energy grid, once the largest in the world, may help Russia to restore at least some economic control over its former empire. □

Zhores A. Medvedev is at 4 Osborn Gardens, Mill Hill, London NW7 1DY, UK.

Reaching for the stars

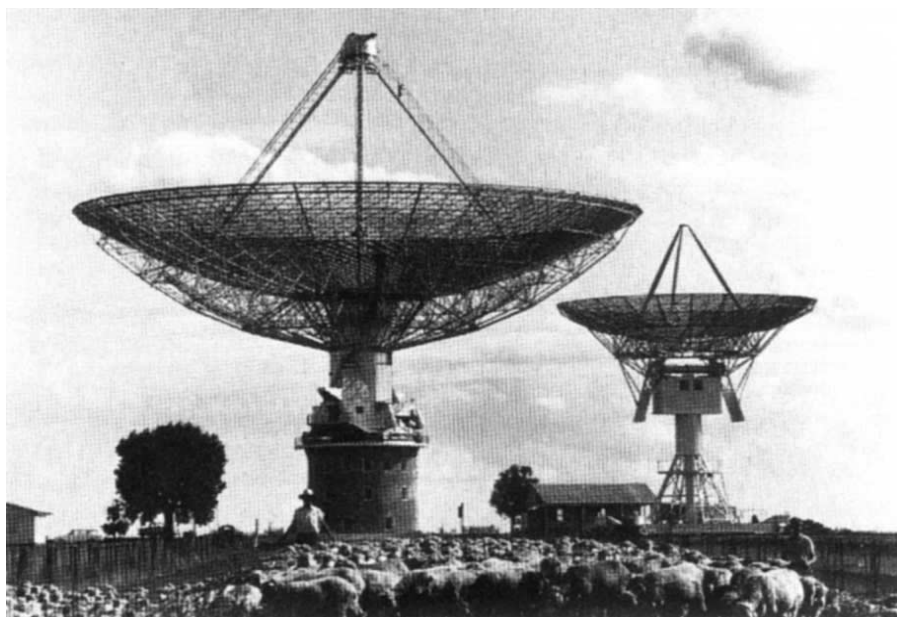
R. Hanbury Brown

Beyond Southern Skies: Radio Astronomy and the Parkes Telescope. By Peter Robertson. *Cambridge University Press: 1993. Pp. 357. £40, \$75.*

OF the many delightful quotations in Peter Robertson's book, the one I like best tells us that "in creating the Universe God perversely located all the most interesting regions of our galaxy in the southern hemisphere but all the astronomers in the north". He certainly

explains how the 21-cm line from interstellar hydrogen was used to map the spiral arms of our Galaxy; he also describes the development of the solar radio spectrograph by Paul Wild and of the high-resolution 'cross' antennae by Bernard Mills and Wilbur Christiansen. I particularly enjoyed reading about John Bolton's early work on radio stars at Dover Heights, and the very fair account of the heated cosmological argument between Martin Ryle and Mills over the number of radio sources. Every one of these topics is skillfully presented in its proper worldwide context.

To be sure, there are a few mistakes in this richly detailed background. I will mention two, neither of which affects the main story. First, in the brief sketch of



From *Beyond Southern Skies*

Big dish, little dish — a 59-ft telescope was installed at Parkes in 1963 to form an interferometer with the main dish. The innovation brought only mixed success, and the small telescope has now fallen into disrepair.

didn't put all the radioastronomers in the north: between 1946 and 1960, which Robertson calls "the golden age" of Australian radioastronomy, the group at the Radiophysics Laboratory in Sydney was larger than any in the United Kingdom, the Netherlands or the United States and was remarkably productive. Researchers had the southern sky all to themselves and radioastronomy was not yet a 'big science': one could still make an important discovery with simple equipment that one had built oneself.

Robertson begins by taking us through this golden age, dedicating the story to Joe Pawsey, to whose wise and expert leadership the Radiophysics Laboratory owed much of its early success. With an agreeable mixture of personal and scientific detail, Robertson tells us about the early Australian work on radio emission from the Sun, the planets and the mysterious radio 'stars' (point sources) and

the early history of radar in the United Kingdom, the earliest work on the detection of aircraft was done not at Bawdsey but at Orford; Bowen's middle name was not Gordon, but George; and as a member of the Airborne Radar Group, I can assure the author that Bowen did not develop airborne radar "single-handedly", nor did we develop the plan-position indicator. And second, the story of how Harry Messel came to create the astronomy department at the University of Sydney is in the wrong order; in fact the tragic death of Colin Gum occurred before, not after, the department was established.

The 210-ft telescope at Parkes was completed in 1961, some four years after the 250-ft Lovell telescope at Jodrell Bank, and in the second part of the book we are told how it was funded and built. For most people such details might be rather boring, but Robertson makes the

narrative come alive through the personalities. He dedicates this part of the story to Taffy Bowen, who was the prime mover. With the help of Fred (later Sir Frederick) White, Bowen managed to keep the project out of the clutches of the Australian National University (at one stage he nearly took the whole setup to the United States), and he used friends in high places to raise large sums of money from the Carnegie and Rockefeller Foundations. We are left in no doubt that fund-raising, and knowing the right people, are essential skills in scientific research.

The account of the design contract with Freeman Fox in London and the subsequent efforts to get the construction done by companies in the United Kingdom presents a depressing picture of British industry as seen from abroad. The presence of Barnes Wallis as a consultant guaranteed an innovative approach, but it wasn't until the tendering was opened to competition outside the United Kingdom that things really began to move. In the end, the main contractor was the West German company Maschinenfabrik Augsburg Nurnberg, which completed the job in roughly two years. The building of the great dish was to split the staff of the Radiophysics Laboratory into two groups, one for and the other against its construction. This division eventually led to the departure of some of the laboratory's senior staff, a loss that became the universities' gain.

The last section of the book is devoted to a description of what has been achieved with the telescope and is dedicated to John Bolton, who has been largely responsible for the dish's successful operation. Faced with more than a thousand scientific papers, the author has chosen to give us the highlights: the measurement of the precise position of the source 3C273 by lunar occultation, which led to the discovery of quasars; the fine work on mapping the southern sky; the detection of the presence of interstellar molecules; and the search for pulsars. Finally, we learn how the telescope has been used to track spacecraft — there is an exciting account of how it came to be the main receiving station for the first moonwalk.

We should welcome this handsome book as a valuable history of Australian radioastronomy; it is well researched, easy to read and generously illustrated. It left me wishing that someone would write something equally attractive about the work in the United Kingdom — golden ages in science are rare and should be recorded. □

R. Hanbury Brown is at White Cottage, Penton Mewsey, Andover, Hampshire SP11 0RQ, UK.

Chains of being

Robert Foley

Self-Made Man: And His Undoing. By Jonathan Kingdon. *Simon and Schuster: 1993. Pp. 369. £25, \$30.*

BOOKS on human evolution fall along a spectrum from the weird to the technical. The weird genre is one of the richest in science, including such gems as Von Daniken's *Chariots of the Gods* (extra-

though, is the last part of the story, the origins of modern humans and the subsequent diversification into geographically variable populations. Kingdon accepts, probably too uncritically, that modern humans originated in Africa about 200,000 years ago and then dispersed across the globe. His main focus is what this dispersal tells us about human variation and evolution. As those who are familiar with Kingdon's other books on East African mammals would expect, the book is copiously illustrated with his own very distinctive line drawings.

Two strong themes run through the book. The first of these derives from Kingdon's background in biogeography. People often think that evolution is a process that occurs solely through time. It is extraordinarily refreshing to see Kingdon's emphasis on the geographical processes and patterns that must underlie any transformation through time. Here we have, albeit speculatively (and more baldly than most specialists would dare), the distributions and dispersals of archaic populations and so a geographical history of human populations throughout the world.

The second theme is set by the title — the self-made man — which harks back to Gordon Childe's *Man Makes Himself* (1936). Childe had in mind that human evolution and development were not just passive responses to an external environment, but were the result of hu-



From *Self-Made Man*

Time's arrow — an example of Kingdon's distinctive art.

manity's own efforts. Human evolution was progressive and driven by intellectual ingenuity as societies invented their own adaptive solutions. Kingdon's view is very different. Humans are not self-made through any inner drive to progress, but through the ordinary processes of adaptation. The do-it-yourself element comes from the fact that the capacity of human populations to use technology to adapt gives each population a unique set of circumstances, which in turn shapes both cultural development and biological characteristics. Variation in skin colour, for example, reflects not only levels of ultraviolet radiation, but also the activities of the people involved; in the same radiant environments, seafarers are darker than terrestrial people because their dependence on tidal patterns constrains them from adopting the shade-seeking behaviour of their land

terrestrial astronauts seeding the Earth with intelligent life) and Branko Bokun's *Man the Fallen Ape* (humans are the result of interbreeding between chimpanzees and orang-utans and have evolved from the laziest apes). The technical are the lifeblood of palaeoanthropology, but seldom make a thrilling read. In the middle of the spectrum are those rare books that are rooted in some sort of evidence while looking at humans as the fascinating evolutionary product that they are. Jonathan Kingdon's *Self-Made Man* is one such book.

In a readable if somewhat diffuse and meandering style, Kingdon explores the history of our species. He starts by looking at longer-term human evolution over the past five or more million years, showing rightly that this is not a simple progressive ladder but a series of adaptive radiations. The central interest,

dwelling counterparts. So technology that builds the boats leads directly to biological adaptation and evolution.

To talk of 'self-made man' in this way is perhaps misleading if it refers to any unique evolutionary mechanisms relating to technology, as Kingdon seems to suggest. All evolutionary change is essentially the result of a tight feedback between the behaviour of the organism and the environments in which it finds itself. What is unique about humans is the rate at which this has occurred and the consequences for other species.

Kingdon's ideas may seem like a recipe for environmental determinism, but the link drawn between a recent origin for humans and the fundamentally local nature of adaptation means that his book is also a powerful indictment of any attempt to rank human populations in terms of progress or to see the evolution of human races as anything other

than short-term responses to particular problems. Human diversity is the product of this process endlessly repeated and re-sorted over tens of thousands of years. So the classic typological races disappear in the scientific reality of time, environment and geography.

Books on human evolution are often written by those who are not specialists in the subject and the results can be catastrophic or embarrassing. But Kingdon's knowledge of the diversity of African animal life and environments enables him to cast a light on human evolution that many specialists would do well to consider seriously (despite reservations about details) and that many others will simply enjoy. □

Robert Foley is in the Department of Biological Anthropology, University of Cambridge, Downing Street, Cambridge CB2 3DZ, UK.

What the British do best

Kevin Burke

Encyclopedia of the Solid Earth Sciences. Editor-in-chief Philip Kearey. Blackwell Scientific: 1993. Pp. 713. £99.50, \$179.95.

THIS is a good book. It is comprehensive and technical enough to meet the needs of professionals and is easy to use.

New to me is the mixing of three styles of article: short definitions, 200-word entries and extended essays of up to 1,500 words. There are 2,700 entries, but the editors emphasize the importance of their alphabetical index with three times as many terms covering key words from the articles and synonyms. The intention is that indexing and cross-referencing should ease the pursuit of a topic in depth. I pursued 'perovskite'. The primary entry was a compositional definition, but two of seven further entries did take me to the deep mantle. There is no text entry for 'dinosaur', but the index gives 11 references that together cover these creatures thoroughly, although some reveal the British distaste for an impact as the cause of the Cretaceous/Tertiary extinction.

Nearly all the 65 contributors work in the United Kingdom, a bias that has slightly influenced the selection of topics and viewpoints. I do not think that "valley bulging" would have had an entry in a similar work from any other country. British strengths emerge. Structural geology is covered well and, among applied topics, so too is engineering geology. Mention of global change is hard to discern, but there is an excellent essay on the history of the Earth sci-

ences. References to books and articles are an asset. Entries are initialised so that it is usually clear who wrote what. Some articles, however, are misleading: for instance, in several places it is indicated that continental material has not been recycled into the mantle.

The illustrations, drawn specially or adapted, are large, clearly labelled and handsome. The conodont animal looks at ease swimming alongside a collection of her teeth, but the accompanying article does not refer to the identification of dentine as a strong indicator of vertebrate affinity. The few photographs are uneven. A Landsat image of Tibet is magnificent but the cliffs of Moher are as black as the O'Loughlins who ruled them. And the many short definitions for minerals seem out of place in an encyclopaedia (I prefer my American Geological Institute glossary, in which I can find 'burkeite').

Some hyperbole on the dust jacket says that the volume is "essential for professional earth scientists". I would not go so far, but this is a good place to start finding out about unfamiliar topics. Students with limited time to complete an essay would do well to begin here, although they are likely to get a British perspective and, inevitably, some less than comprehensive coverage. As a final thought, my enjoyment of the longer essays made me regret that we still do not have an *Oxford Companion to the Earth Sciences*. □

Kevin Burke is in the Department of Geosciences, University of Houston, Houston, Texas 77204-5503, USA.

Untangling data and theory

Deborah M. Gordon

Spiders in Ecological Webs. By D. H. Wise. Cambridge University Press: 1993. Pp. 328. £45, \$79.95.

COMMUNITY ecology is struggling to bring data and theory together. Is it possible to establish any general principles of ecology? Do we need more data or better theory? This book contributes to the debate by looking at existing data on spiders. Wise considers some basic problems of community ecology: how important competition is; whether niche partitioning structures communities; how much can be explained by food webs; and how to determine the effects of one population on the dynamics of another. His discussion is based mainly on the results of field experiments, and he concludes that spiders are limited in numbers by food supplies, yet do not compete with each other for food or limit the population sizes of their prey. The tone of the book is upbeat, with the author arguing that more data of the right sort will resolve the outstanding questions of community ecology.

The empirical literature is carefully reviewed, with an emphasis on the evaluation of field methods, data analysis and experimental design. At least to someone outside the field of spider ecology, Wise's discussion of his colleagues' work seems fair and even-handed. The writing is straightforward and generally very accessible, although it is sometimes prone to metaphorical entanglements in discussions of the big picture. Spider lovers will enjoy an amazing abundance of spiderly puns; the title is only the beginning.

The book will be useful reading for anyone seeking to bridge the gap between data and theory in ecology. It should make an interesting text for a graduate-level course in community ecology, help students to plan research and set any ecologist thinking about possible generalizations. □

Deborah M. Gordon is in the Department of Biological Sciences, Stanford University, Stanford, California 94305-5020, USA.

New Journals issue

This year, *Nature's* annual New Journals review supplement will appear in the issue of 7 October. Publishers and learned societies are invited to submit journals for review. For further information please see last week's issue or telephone Peter Tallack on 071-836-6633 (011-44-71-836-6633 from the United States), extension 2414.

The pathogenesis of atherosclerosis: a perspective for the 1990s

Russell Ross

Atherosclerosis, the principal cause of heart attack, stroke and gangrene of the extremities, is responsible for 50% of all mortality in the USA, Europe and Japan. The lesions result from an excessive, inflammatory-fibroproliferative response to various forms of insult to the endothelium and smooth muscle of the artery wall. A large number of growth factors, cytokines and vasoregulatory molecules participate in this process. Our ability to control the expression of genes encoding these molecules and to target specific cell types provides opportunities to develop new diagnostic and therapeutic agents to induce the regression of the lesions and, possibly, to prevent their formation.

ATHEROSCLEROSIS is not merely a disease in its own right, but a process that is the principal contributor to the pathogenesis of myocardial and cerebral infarction, gangrene and loss of function in the extremities. The process, in normal circumstances a protective response to insults to the endothelium and smooth muscle cells of the wall of the artery, consists of the formation of fibrofatty and fibrous lesions, preceded and accompanied by inflammation. The advanced lesions of atherosclerosis, which, when excessive, become the disease and which may occlude the artery concerned, result from an excessive inflammatory-fibroproliferative response to numerous different forms of insult.

This view of the atherosclerotic plaque, as the response to insult, formed the basis of my previous review in 1986¹. Since then, there has been remarkable progress in the identification of different cell types and the various molecules that are involved in atherogenesis. At the same time, clinical investigations have confirmed the observation from animal experiments that the lesions of atherosclerosis can be reversed. What follows is a review of what is now known of the cells and molecules involved

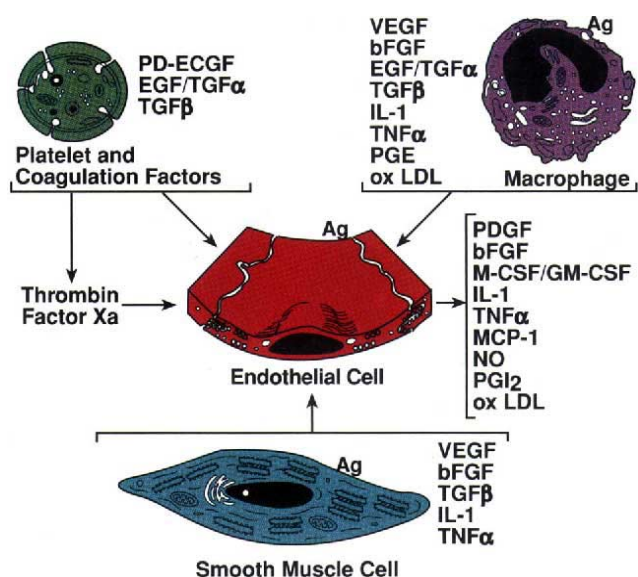
and of the potential for applying what is known of their interaction to the prevention and the treatment of this disease.

Lesions of atherosclerosis

The earliest recognizable lesion of atherosclerosis is the so-called 'fatty streak', an aggregation of lipid-rich macrophages and T lymphocytes within the innermost layer of the artery wall, the intima. The ubiquity of the atherosclerotic process is attested to by the finding of fatty streaks in the coronary arteries of half of the autopsy specimens from children aged 10 to 14 (refs 2-4). Animal observations have shown that fatty streaks precede the development of intermediate lesions⁵⁻¹⁰, which are composed of layers of macrophages and smooth muscle cells and, in turn, develop into the more advanced, complex, occlusive lesions called fibrous plaques. The fibrous plaques increase in size and, by projecting into the arterial lumen, may impede the flow of blood. They are covered by a dense cap of connective tissue with embedded smooth muscle cells that usually overlays a core of lipid and necrotic debris. The fibrous plaques contain

BOX 1 Endothelium in atherosclerosis

In the genesis of the lesions of atherosclerosis, the endothelium can interact with macrophages, platelets and smooth muscle, as well as T lymphocytes. The principal products of platelets, macrophages and smooth muscle that may affect the endothelium are shown in this diagram. Endothelial mitogens produced by macrophages include vascular endothelial growth factor (VEGF), FGF and transforming growth factor α (TGF α). TGF β , as well as interleukin-1 (IL-1) and TNF α are all capable of inhibiting endothelial proliferation. Furthermore, these three molecules can induce secondary gene expression of other growth-regulatory molecules, such as PDGF, by the endothelium. TGF β can also induce endothelial synthesis and secretion of connective tissue matrix. Oxidized LDL produced by endothelial cells or macrophages can have a profound effect in furthering the injury to the endothelial cells. Similarly, platelets can provide not only TGF α and TGF β but a potent endothelial mitogen, platelet-derived endothelial cell growth factor (PD-ECGF). Thrombin and Factor Xa from plasma may also stimulate the endothelium to a procoagulant state. Several of the molecules that can be formed by macrophages and platelets can also be generated by smooth muscle cells in the artery wall or in the lesions of atherosclerosis that underlie the endothelium. In turn, the endothelial cells can express genes for and synthesize a number of growth-regulatory molecules (see box on figure), including molecules that can induce connective tissue proliferation, such as PDGF, bFGF and TGF β , or IL-1 and TNF α , which can induce secondary gene expression for PDGF in endothelium and smooth muscle. The endothelial cells can also be a principal source of mitogenic and activating factors for the underlying macrophages through M-CSF, GM-CSF or oxLDL. Similarly, the endothelial cells can participate in leukocyte chemotaxis by providing the potent chemotactic factors, oxLDL and MCP-1. Through their capacity to form NO and PGI₂, endothelial cells can profoundly



affect the state of dilation or constriction of the artery. The endothelial cells represent the first potential site of oxidation of LDL as it transports the LDL into the artery wall. For colour coding see Fig. 1. □

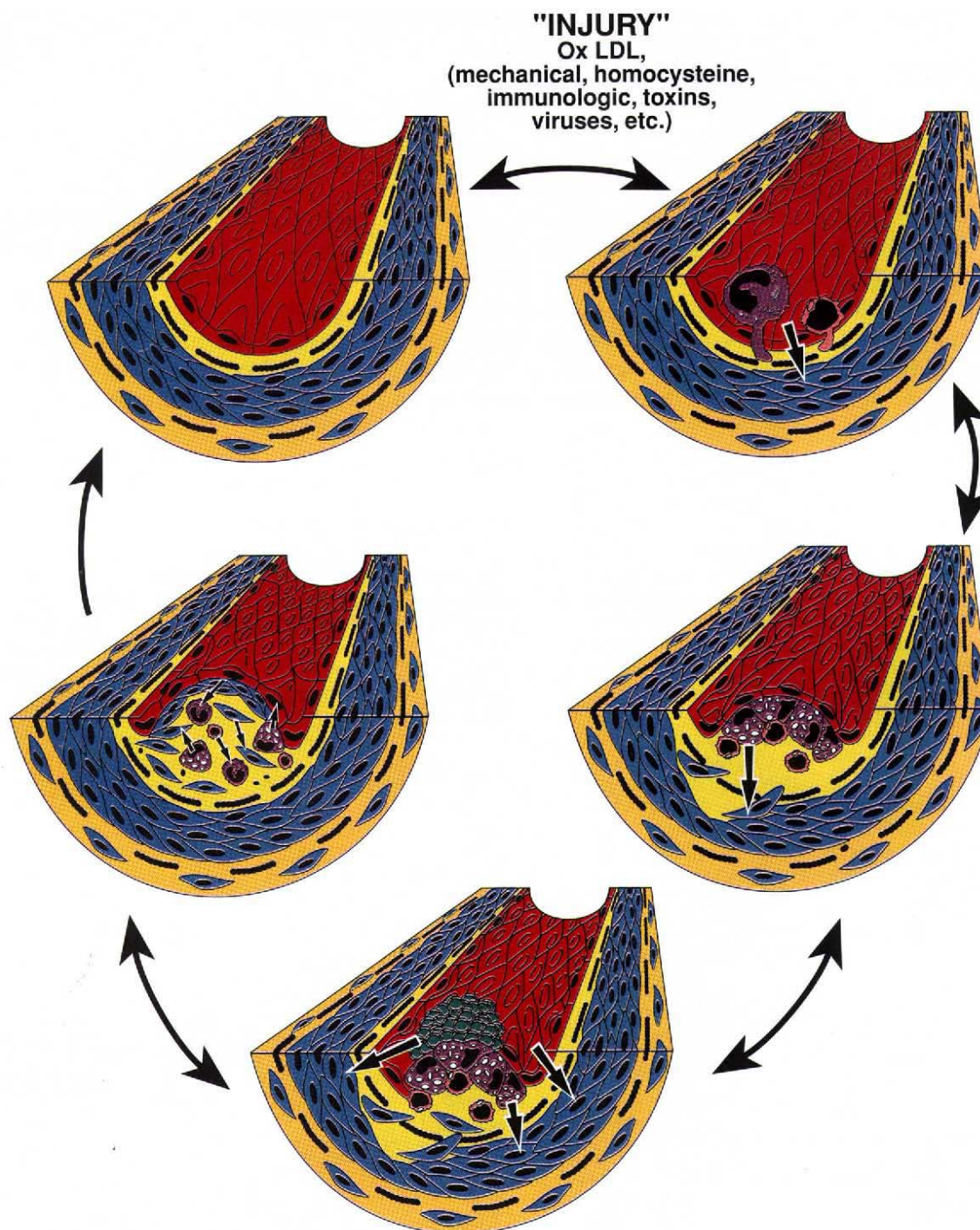


FIG. 1 The response-to-injury hypothesis of atherosclerosis. Several different sources of injury to the endothelium can lead to endothelial cell dysfunction. One of the parameters associated with endothelial cell dysfunction that results from exposure to agents, such as oxLDL, is increased adherence of monocytes/macrophages and T lymphocytes. These cells then migrate between the endothelium and localize subendothelially. The macrophages become large foam cells because of lipid accumulation and, with the T cells and smooth muscle, form a fatty streak. The fatty streak can then progress to an intermediate, fibrofatty lesion and ultimately to a fibrous plaque. As the lesions accumulate more cells, and the macrophages scavenge the lipid, some of the lipid-laden macrophages may emigrate back into the bloodstream by pushing apart the endothelial cells. On doing so, those at sites such as branches and bifurcations, where blood flow is irregular with eddy currents and back currents, may become thrombogenic sites that lead to formation of platelet mural thrombi. Such thrombi can

release many potent growth-regulatory molecules from the platelets that can join with those released by the activated macrophages and possibly by lesion smooth muscle cells into the artery wall. Platelet thrombi can also form at sites where endothelial dysfunction may have occurred. Ultimately, the formation and release of numerous growth-regulatory molecules and cytokines from a network established between cells in the lesion, consisting of activated macrophages, smooth muscle, T cells and endothelium, lead to progression of the lesions of atherosclerosis to a fibrous plaque or advanced, complicated lesion. Each of the stages of lesion formation is potentially reversible. Thus, lesion regression can occur if the injurious agents are removed, or when protective factors intervene to reverse the inflammatory and fibroproliferative processes. Cells have been colour coded in this figure and in the boxes: smooth muscle, blue; endothelium, red; macrophage, violet; T cell, pink; platelet, green.

monocyte-derived macrophages, smooth muscle cells and T lymphocytes, many of which are activated, as evidenced by HLA-DR expression¹¹⁻¹⁵. Recent data have shown that most of the sudden deaths from myocardial infarcts are due to ruptures or fissures, particularly in the margins of the fibrous cap where there are more macrophages, resulting in haemorrhage into the plaque, thrombosis and occlusion of the artery¹⁶.

Response to injury and atherogenesis

The hypothesis that an injury to the endothelium might precipitate the atherosclerotic process, a notion rooted in earlier hypotheses of arterial injury^{17,18}, was formally advanced in 1973¹⁹. It has been continually modified²⁰⁻²², particularly in 1986¹, to take account of accumulating observations. Manifestations of the dysfunction of the endothelium caused by injury, notably at branch points in the arterial tree, include increased trapping of lipoprotein in the artery²³ and the appearance of specific adhesive glycoproteins on the surfaces of the endothelial cells. Monocytes and T lymphocytes attach to them and migrate between the endothelial cells under the influence of growth-regulatory molecules and chemoattractants released both by the altered endothelium, its adherent leukocytes²⁴⁻²⁷, and possibly by underlying smooth muscle cells.

As the process continues, migrating cells reach further beneath the arterial surface, where the monocytes become macrophages, accumulate lipid, become foam cells and, together with the accompanying lymphocytes, become the fatty streak⁵⁻¹⁰. These often form at sites of pre-existing collections of intimal smooth muscle²⁸. Thereafter, continued cell influx and proliferation lead to the more advanced lesions, distinguished by their fibrous character, and ultimately to the fibrous plaque (Fig. 1).

Studies of animals with artificially induced hypercholesterolaemia^{5-10,23,28,29} have confirmed that three processes are involved in the formation of atherosclerotic lesions: (1) the proliferation of smooth muscle cells, macrophages and possibly lymphocytes; (2) the formation by smooth muscle cells of a connective tissue matrix comprising elastic fibre proteins, collagen and proteoglycans; and (3) the accumulation of lipid and mostly free and esterified cholesterol in the surrounding matrix and the associated cells.

The cellular events that occur during progression of lesions in hypercholesterolaemic animals are almost exactly mirrored by those observed in human atherosclerotic coronary arteries in hearts removed in transplant operations³⁰. Figure 2*a-d* compares the responses in nonhuman primate (right) and human (left) arteries. Shortly after the induction of hypercholesterolaemia, monocytes and lymphocytes adhere in clusters to the endothelium (Fig. 2*a*), migrate over its surface, and reach the subendothelial intima by penetrating at endothelial cell junctions (Fig. 2*b*). There the monocytes are converted to macrophages, ingest lipids, become foam cells and so on. The fatty streak expands and, as a consequence, the luminal surface of the artery becomes convoluted (Fig. 2*c*).

Progression of atherosclerotic lesions is thus marked by the accumulation of alternating layers of smooth muscle cells and lipid-laden macrophages. In monkeys and rabbits, as well as in humans, the arteries contain sites with retracted endothelial cells, which expose underlying lipid-filled macrophages, which, in turn, provide sites for platelet interactions leading to the formation of mural thrombi (Fig. 2*d*). Many of these exposed macrophages seem, as in other inflammatory responses, to be attempts by the macrophages to remove the accumulating lipid by circulating back to the lung, liver, spleen and lymph nodes⁶. Advanced atherosclerotic lesions may develop at the same sites, particularly branches and bifurcations, where the mural thrombi occur.

The formation of fibrous lesions in response to injury is not in itself remarkable; wound healing follows such a course, for example. But the response to arterial injury differs from that to the injury of most other tissues and organs in two respects. First,

the principal source of connective tissue in the arterial wall is the smooth muscle cell. Second, the sources of arterial injury (hypercholesterolaemia, hypertension, cigarette smoking, diabetes, obesity and so on) are likely to be chronic, so that the progression from fatty streak to fibrous plaque is unlikely to be interrupted or, at best, may be episodic.

Molecules

Growth factors, cytokines and other chemicals, including lipids and small molecules such as nitric oxide (NO), induce and regulate numerous critical cell functions. During the process of atherogenesis, they may act in cell recruitment and migration, cell proliferation and the control of lipid and protein synthesis (including extracellular matrix proteins). Furthermore, the same modifiers are implicated in vascular events, such as vasodilation, vasoconstriction and coagulation. Elucidation of their roles in this disease process provides opportunities to use them as potential starting points for therapeutic intervention.

Growth factors and cytokines. The terms growth factor and cytokine are often used interchangeably, although the cytokines were originally considered to be the mediators implicated in immunity and inflammation, and growth factors are involved in proliferation and chemotaxis of cells in organs and tissues. With respect to atherosclerosis, these two roles are closely inter-related, with the same mediator sometimes acting as an inflammatory regulator and as a growth regulator, depending on the cell that is the target. These complex interactions are illustrated in Boxes 1-3.

The growth-regulatory molecules can induce multiple and, in some instances, apparently divergent effects. They can stimulate cell proliferation, can in some cases inhibit proliferation, and many of the proliferative agents can act as chemoattractants. Several of the molecules potentially important in cell proliferation include platelet-derived growth factor (PDGF)³¹⁻³⁴, basic fibroblast growth factor (bFGF)³⁵⁻³⁷, heparin-binding epidermal growth factor-like growth factor (HB-EGF)³⁸, insulin-like growth factor-1 (IGF-1)³⁹, interleukin-1 (IL-1)^{40,41}, tumour necrosis factor α (TNF α)⁴² and transforming growth factor β (TGF β)⁴³⁻⁴⁹. These can all induce smooth muscle proliferation and are generally not expressed in the normal artery, whereas they are upregulated in lesions of atherosclerosis.

Several growth factors that are mitogens are also chemoattractants (Box 2). Chemotaxis is a critical event in the development of the lesions of atherosclerosis and in the restenosis that often occurs after surgical intervention and angioplasty. Thus, chemotaxis is necessary to bring leukocytes into the artery wall and, at some sites, smooth muscle cells from the media into the intima of the artery. Colony-stimulating factors (CSFs)⁵⁰, monocyte chemoattractant protein-1 (MCP-1)^{51,52}, oxidized low-density lipoprotein (oxLDL)⁵³⁻⁵⁵, and TGF β ⁴⁴ can each induce monocyte chemotaxis and endothelial transmigration, whereas PDGF^{32,56} and IGF-1 can induce smooth muscle chemotaxis. Basic FGF³⁷, which has no signal sequence and is present in the cytosol of most cells, is also found in basement membranes and could be released as a result of cell injury. It is a potent mitogen and chemoattractant for endothelium and a mitogen for smooth muscle.

The cytokines, IL-1, TNF α , interferon γ (IFN γ)⁵⁷, and IL-2, together with the CSFs, are modulators of the inflammatory response that occurs once the endothelium has been exposed to injurious agents. The nature of any immune response that is induced locally during atherogenesis is not understood, but presumably these same cytokines can also play a role as immunological mediators as well (see below).

It is probable that none of these factors works alone in the process of atherogenesis. Through a network of cellular interactions, the release of one molecule can lead to expression of a second molecule in a target cell that can then either stimulate its neighbours in a paracrine way, or itself in an autocrine way. Examples of this type of interaction include release of the

cytokines, IL-1 or TNF α , by activated lesion macrophages exposed to agonists, such as oxLDL^{54,58}. In culture, either of these macrophage products, or another molecule, TGF β , will induce PDGF-A gene expression in smooth muscle cells⁴⁹ (Box 4). *In vivo*, PDGF-AA secreted by such activated smooth muscle cells could then stimulate their neighbours or themselves⁵⁹.

Similarly, when endothelial cells in culture are exposed to agonists, including IL-1⁴¹, TNF α ⁶⁰ or TGF β ⁴⁹, they express the gene for and secrete PDGF-BB. Thus if these cytokines were released *in vivo* by activated underlying macrophages, PDGF-BB secretion by the adjacent endothelium could in turn stimulate underlying smooth muscle cells to migrate or replicate (Box 4). Consequently, cytokines such as these, as well as several other molecules, may effect a chain reaction of sorts that could lead to amplification of cell accumulation. As a potential balance to these mitogens, several factors can inhibit cell proliferation. These include TGF β ^{46,48}, IL-1⁴⁰ and IFN γ ⁵⁷, all of which can inhibit either endothelial or smooth muscle proliferation. An apparently dichotomous effect occurs in the presence of excess IL-1, TNF α or excess TGF β . When each of these is present in high doses, it can inhibit the proliferative effect of the secreted PDGF-AA that it has induced by downregulating the relevant PDGF receptor molecules⁴⁹. This appears to provide a self-regulatory means of preventing overstimulation by the PDGF it initially induces.

Support for the relevance of these observations in cell culture has come from data demonstrating PDGF gene expression is increased in smooth muscle cells and in adjacent macrophages in human lesions of atherosclerosis^{61,62}. Specimens obtained by endarterectomy of the carotid artery contain increased messenger RNA for PDGF B-chain and for *c-fms*, the proto-oncogene that encodes the M-CSF receptor, prominent on macrophages and recently observed on smooth muscle cells⁶³. Similarly, *in situ* hybridization and immunocytochemistry have demonstrated localization of PDGF A-chain mRNA in 'mesenchymal cells'⁶¹ and PDGF B-chain protein in lesion macrophages⁶², respectively. In fact, lesions at all stages of development from both humans and nonhuman primates contain PDGF B-chain (presumably PDGF-BB) in Golgi/endoplasmic reticulum-type distribution in as many as 25% of the macrophages in these lesions⁶². The capacity to respond to this mitogen is associated with increased numbers of PDGF β -receptors on adjacent intimal smooth muscle cells in the same lesions⁶⁴.

Regulation

Several methods to control exposure of the cells to these growth-regulatory molecules and cytokines have evolved, so that the molecules can be prevented from acting or may reside in the tissues for long periods of time. For example, most of the growth-regulatory molecules have a relatively short half-life in the circulation, mostly because the presence of binding proteins, such as α 2 macroglobulin, a principal component of plasma that can bind to many of these molecules, including PDGF, TGF β , IL-1, IL-2, IL-6 and bFGF⁶⁵. Under these circumstances, substances such as α 2 macroglobulin may prevent the action of these molecules. In addition, a number of molecules in connective tissue can affect the storage and activity of growth factors. For example, many glycosaminoglycans, in particular heparin, a product of endothelium and smooth muscle, can inhibit both migration and proliferation of smooth muscle cells^{37,65}. Heparin can bind to and sequester PDGF and bFGF, can release TGF β from a carrier protein, and can modulate the phenotype of the smooth muscle cells^{47,65,66}. Other matrix molecules, such as fibronectin and laminin, can have profound effects on smooth muscle cell phenotype and on the capacity of the cells to respond to mitogens and chemoattractants and, importantly, on the ability of the cells to adhere to the matrix and to other cells. These substances can thus alter homeostasis by modifying cytoskeletal elements and cell-cell surface

adhesion receptors, such as the integrins. This in turn can have profound effects on the cellular response to growth-regulatory molecules and cytokines because the disruption of integrins, particularly those associated with cell-matrix interactions, will lead to altered cell regulatory activities, including transmission of intracellular signals, such as tyrosine phosphorylation, pH alterations in the cytosol, and association of subplasmalemmal agents with the plasma membrane during the process of signalling.

During the most advanced stages of atherogenesis as the lesions become thicker, fibrous plaques become vascularized and contain large numbers of capillary and venule-like channels. Formation of such secondary vascularization may be due to release of substances such as bFGF or other angiogenic molecules^{36,37}. These channels could play a critical role in the terminal phases of atherosclerosis when alterations in the structure of the plaques, such as fissuring, cracking, or ulceration, lead to haemorrhage from the lumen or from these small vessels, and set the stage for the formation of a terminal thrombus¹⁶.

As part of the normal repair process, accumulation of the different matrix constituents will alter the deformability of the lesion and thus of the artery wall. One of the most ubiquitous and potent factors potentially involved in matrix formation is TGF β ^{44,45}. Not only is it a potent inhibitor of mitogenesis, but it is the most potent stimulator of connective tissue formation known. It can induce gene expression for fibronectin, collagen types I, III, IV and V, as well as numerous glycosaminoglycan and proteoglycan moieties and elastic fibre proteins. In most circumstances, TGF β is present in a latent form that requires exposure to acidic conditions or specific enzymes, such as plasmin or trypsin-like proteases, to cleave an accompanying peptide and thus release active TGF β ^{44,45}. TGF β seems to inhibit cell proliferation by suppression of transcription of the proto-oncogene *c-myc*, as well as by inducing increased deposition of connective tissues by the smooth muscle cells^{43,44,46}.

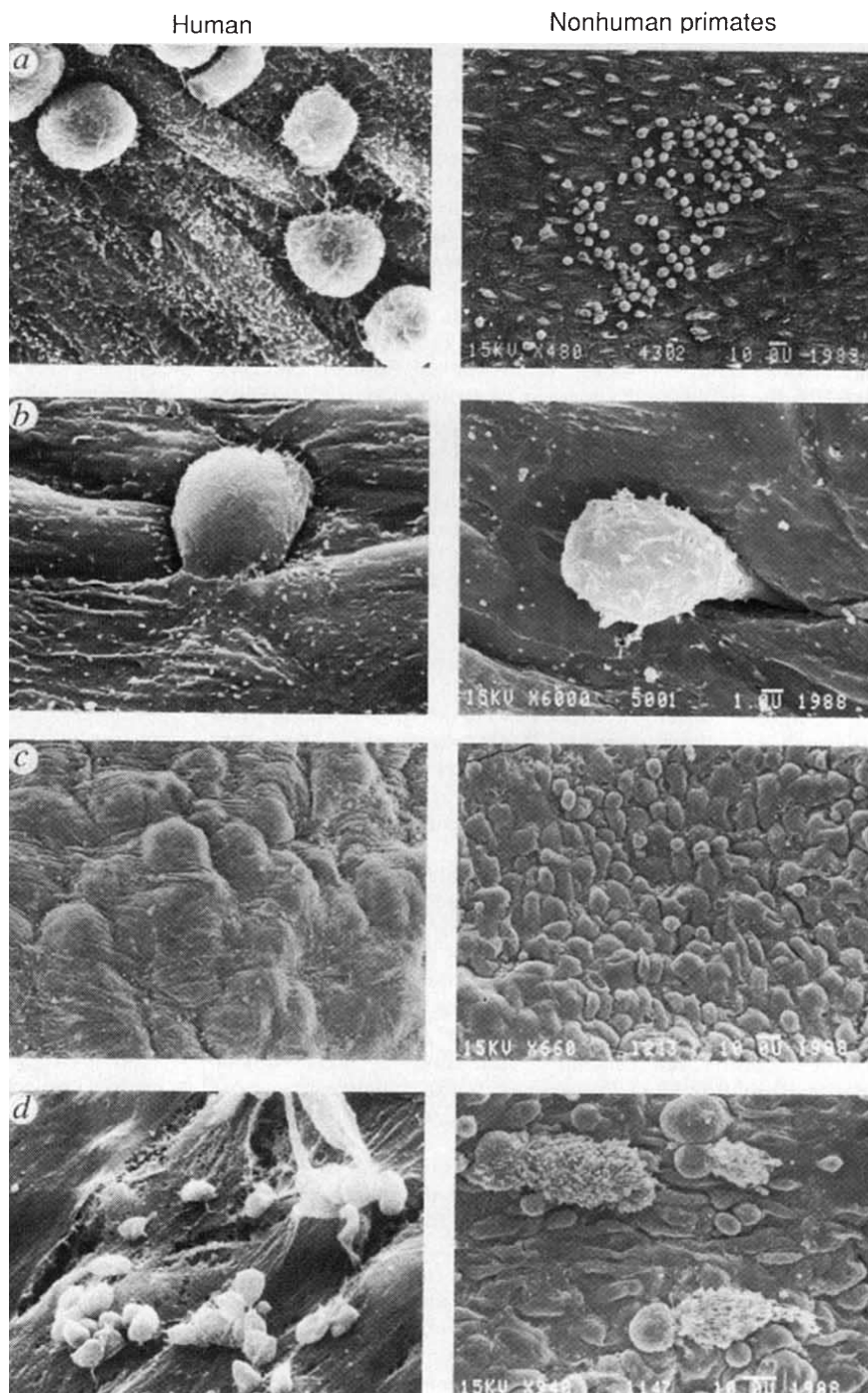
Not only is smooth muscle replication an important feature of atherogenesis, but proliferating macrophages seem to be important in this process as well^{67,68}. It is not clear whether the monocytes have entered into S phase in the marrow, in the plasma or after entering the artery wall. Consequently, the role of increased gene expression and formation of M- and GM-CSF in this process needs to be pursued further. Both human and rabbit lesions contain immunohistochemically identifiable M-CSF, as well as increased M-CSF mRNA⁶⁹. Curiously, increased M-CSF in the plasma can lead to lowering of plasma cholesterol levels⁷⁰. One possible explanation is that M-CSF upregulates scavenger receptors and possibly clearance of lipoproteins.

Cellular interactions

Endothelium. Central to the response-to-injury hypothesis^{1,19-22} is the proposal that the different risk factors somehow lead to endothelial dysfunction, which can elicit a series of cellular interactions that culminate in the lesions of atherosclerosis. The initial injurious events do not necessarily lead to endothelial denudation. In fact, with better modes of tissue preservation, it is clear that the early lesions develop at sites of morphologically intact endothelium⁵⁻¹⁰.

What then is the nature of the dysfunctional changes in the endothelium? Endothelial cells play numerous physiological roles⁷¹ including: (1) provision of a nonthrombogenic surface; (2) a permeability barrier through which there is exchange and active transport of substances into the artery wall; (3) maintenance of vascular tone by release of small molecules such as NO, prostacyclin (PGI₂), and endothelin (ET) that modulate vasodilation or vasoconstriction, respectively; (4) formation and secretion of growth-regulatory molecules and cytokines; (5) maintenance of the basement membrane collagen and proteoglycans upon which they rest; (6) provision of a nonadherent surface for leukocytes; and (7) ability to modify (oxidize) lipoproteins as they are transported into the artery wall. Changes

FIG. 2 Scanning electron micrographs (SEM) taken of the aorta from hypercholesterolaemic nonhuman primates (right panels) versus coronary arteries obtained from human hearts with advanced atherosclerosis removed for transplant purposes (left panels) that were replaced by healthy hearts. The atherosclerotic hearts were rapidly perfuse fixed and prepared for SEM. *a* (Right), Numerous adherent leukocytes are clustered on the surface of the endothelium in this monkey artery in an area suggesting that there are localized cell surface changes at that site. Increased monocyte and T lymphocyte adherence precedes their entry into the artery wall. (Left), At a somewhat higher magnification, leukocytes from a human coronary artery are similarly attached to the surface of the endothelium. Most of the cells are rounded, although a few of them appear to be in the process of spreading on the surface of the endothelium. *b* (Right), A leukocyte that has begun to penetrate between the surface endothelial cells. The cell has extended a cytoplasmic process that appears to protrude between the endothelial cells as it enters into the subendothelial space of the monkey aorta. (Left), A similar process of leukocyte chemotaxis appears to be taking place in a human carotid artery, in which the leukocyte appears to be burrowing beneath one of the lining endothelial cells. From Davies *et al. Br. Heart J.* **60**, 459–464 (1988); reproduced with permission. *c* (Right), A surface of a fatty streak taken from a hyperlipidaemic monkey. The lobular appearance of the surface of the lesion is due to the large number of accumulated underlying lipid-filled macrophages or foam cells. Monocytes and lymphocytes can be seen to continue to adhere to the surface of the lesion. In this way, their continual entry will lead to further lesion expansion. (Left), A similar lobulation of a fatty streak from a human coronary artery. The underlying foam cells with their lipid droplets can be seen through the stretched, thinned endothelium. The processes in the human appear to be virtually identical to those observed in the nonhuman primate. *d* (Right), Three mural platelet thrombi on the surface of an intermediate lesion of atherosclerosis in a nonhuman primate. When sectioned transversely and examined by light microscopy, some of these can be shown to be located on the surfaces of exposed macrophages, others at sites of endothelial cell separation. (Left), Adherent platelets are present on the surface of a lesion of atherosclerosis from a human coronary artery. In this case, they are associated with an area of endothelial separation and exposure of the underlying connective tissue. In both cases, adherent, aggregated platelets undergo shape changes and characteristically release the numerous growth factors and other molecules in their granules.



in one or more of these properties may represent the earliest manifestations of endothelial dysfunction.

Evidence has accumulated to suggest that oxLDL is a key component in endothelial injury^{53,54,58,72}. Once formed by the endothelium, oxLDL may directly injure the endothelium and play an initial role in the increased adherence and migration of monocytes and T lymphocytes into the subendothelial space (Figs 1, 2*a, b* and Box 1). It can induce formation of adhesive cell-surface glycoproteins, such as VCAM-1 (athero-ELAM), by the endothelium. Several classes of leukocyte-endothelial adhesion molecules are probably involved in atherogenesis^{73–75}.

Once monocytes and lymphocytes enter into the intima of the artery, oxLDL from the endothelium and other substances associated with atherogenesis may participate in activation of the monocytes into macrophages⁷⁶. Uptake of oxLDL by the

macrophages will lead to foam cell formation⁷⁶ and may alter gene expression of many of the growth-regulatory molecules, cytokines, and substances of low molecular mass, noted above. Curiously, additional oxLDL may also form as a result of the action of NO⁷⁷ or the enzyme responsible for leukotriene formation lipoxygenase^{78,79}, both of which can be released by the endothelial cells, presenting a dilemma. On the one hand, formation of NO by endothelium can prevent platelet aggregation and leukocyte adhesion and promote vasodilation^{80,81}; yet, on the other hand, it can promote oxidation of LDL and all of the deleterious consequences of its formation^{77,82}. Vasomotor tone of the artery appears to be controlled by the constant action of NO⁸¹. Inhibition of formation of NO or eicosanoids, such as PGI₂, would permit opposing forces of vasodilation, resulting from vasoconstrictors such as ET, angiotensin II (A-II), or thromboxane A-2, to determine the capacity of the artery to

maintain its lumen in the presence of the changing forces that result from formation and progression of the lesions of atherosclerosis⁸³⁻⁸⁷.

Endothelial cells are relatively unique, because they represent a cell population that grows as a strict monolayer. Cell-matrix integrins as well as intercellular attachments may be critical in controlling the ability of the endothelium to respond to factors released by neighbouring cells. One possible role for these attachments may be associated with contact inhibition of endothelium. When endothelial cells are wounded, contact between the cells is disrupted, and the cells closest to the wound will maintain their remaining attachments and stretch into the wound in attempts to re-establish their lost intercellular connections. If the wound is too large to re-establish contact, the cells closest to the wound will undergo replication. Hypothetically, if this were to occur because of chronic injury at the same site, cells at the wound margin might undergo sufficient doublings so that ultimately they would approach senescence, whereas cells distant from the wound could retain the capacity to replicate but be unable to participate in wound closure because of the inhibitory restrictions of the monolayer. Thus, repeated injury could theoretically result in loss of capacity of the cells at the wound margin to replicate and lead to alterations in the surface of the artery at the injured site.

Endothelial cells are also subject to other forms of injury such as those that can be induced by viruses. A herpes simplex type of viral infection can induce lesions in avian species that appear very much like atherosclerosis⁸⁸⁻⁹⁰. Herpes simplex I can enter endothelial cells through FGF receptors on their surface⁸⁸. Although both herpes and cytomegalovirus have been found in arteries, a role for them has not been demonstrated in human atherosclerosis.

Finally, endothelium plays a critical role in both thrombotic and coagulant activities. Heparan sulphate⁹⁰ and production of NO⁸⁰ and PGI₂⁸³ on endothelial surfaces are all antithrombotic, and endothelium can specifically bind factors that prevent coagulation⁹¹ (Box 1). Additionally, the endothelium can balance interactions between the coagulation and fibrinolytic systems by forming plasminogen activator and urokinase⁹². In lesions of atherosclerosis, the formation of plasminogen activator inhibitor (PAI-1) could shift this balance to a pro-thrombotic state, which would be particularly deleterious in the latter stages of the disease^{93,94}.

Monocytes/macrophages. Macrophages are present in all stages of atherogenesis^{1,11,12}. The normal role of the macrophage is to act not only as an antigen-presenting cell to T lymphocytes, but also as a scavenger cell to remove noxious materials and as a source of growth-regulatory molecules and cytokines. The macrophage is thus the principal, inflammatory mediator of cells in the atheromatous plaque microenvironment.

Macrophages internalize oxLDL through both scavenger receptors⁹⁵ and a putative oxLDL receptor and can oxidize LDL through several pathways, including lipoxygenase enzymes⁹⁶. During this process, fatty acids undergo peroxidation, yielding a variety of short chain aldehydes, ketones and other substances⁹⁸, which can become covalently crosslinked to the apoprotein B moiety of the LDL particle and thus can bind to the scavenger receptor and be taken up by macrophages⁹⁵. If the generation of oxLDL were excessive, this could set up a vicious cycle that could further exacerbate the atherogenic process⁹⁷. The importance of oxLDL in atherosclerosis was first established through the use of the antioxidant, probucol, in studies of a genetic hyperlipidaemic model of atherosclerosis, the Watanabe heritable hyperlipaemic (WHHL) rabbit^{98,99}. As a result, probucol and more recently vitamin E, vitamin C and possibly betacarotene have been suggested as a means of preventing and reducing lesions of atherosclerosis. Recent observations by Sasahara *et al.* (unpublished data) have demonstrated that probucol is equally effective in reducing lesions of atherosclerosis in the nonhuman primate, setting the stage for a clinical

trial of antioxidants in the prevention and treatment of the human disease.

As mentioned above, not only are macrophages important as scavenger cells, but they also play a role in the fibroproliferative process by their capacity to form numerous growth factors (Box 2), in particular PDGF^{62,100} as well as IL-1⁴⁰ and TNF α ¹⁰¹, which can result in secondary PDGF gene expression by smooth muscle cells or endothelium (Box 4). One particular form of atherosclerosis where macrophages and T lymphocytes occur in very large numbers are the lesions associated with cardiac transplant rejection^{15,102}. The extraordinary numbers of macrophages and T cells, compared with those observed in common atherosclerosis, suggest that a localized immune response may have exacerbated macrophage/T-cell interactions and macrophage/smooth muscle interactions leading to pronounced proliferation of all of these cells. A similar role for macrophages also can occur in common atherosclerosis.

Numerous questions remain to be answered in studies of the macrophage in atherogenesis. In particular it will be important to determine whether there are specific subsets of monocytes that enter the lesions from the circulation and whether many of the cells that are found in cell cycle traverse in the lesion represent cells that had already entered the cycle in the marrow before release into the circulation or whether cell cycle traverse is stimulated in the lesions themselves. The critical role the macrophages may play in the inflammatory response in lesion formation suggests that finding means of controlling macrophage participation at all levels of atherogenesis could be critical in modifying lesion progression.

Smooth muscle. There may be smooth muscle diversity in the lesions of atherosclerosis, because smooth muscle cells are locally derived from individual organ parenchyma during embryogenesis, whereas the endothelium appears to come from the embryonic vasculature that invades the organ¹⁰³. Thus smooth muscle cells in different arteries may in fact be different and may respond differently to agonists with which they are locally presented, explaining in part why different arterial beds appear to respond differently to local stimuli associated with the process of atherogenesis¹⁰⁴.

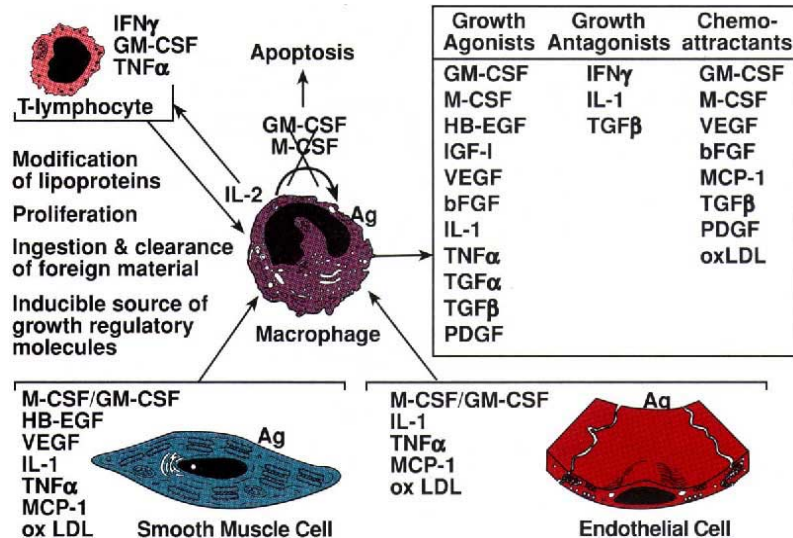
At least two different phenotypes have been described for smooth muscle cells based on the distribution of myosin filaments and the formation of large amounts of secretory protein apparatus, such as rough endoplasmic reticulum and Golgi¹⁰⁵⁻¹⁰⁷. When cells are in a contractile phenotype, they respond to agents that induce either vasoconstriction or vasodilation, such as ET¹⁰⁸, catecholamines¹⁰⁹, A-II¹¹⁰, PGE¹¹¹, PGI₂¹¹², neuropeptides¹¹³, leukotrienes¹¹⁴ and NO¹¹⁵. In contrast, in the synthetic state, they are capable of expressing genes for a number of growth-regulatory molecules and cytokines^{59,116}, can respond to growth factors by expressing appropriate receptors, and can synthesize extracellular matrix (Box 3). Numerous observations suggest that smooth muscle cells in lesions have changed from a contractile to a synthetic phenotype, which could have profound effects on the capacity of the lesion to respond to various agonists¹¹⁶. As shown in Box 3, smooth muscle cells in the synthetic phenotype can respond in an autocrine way to PDGF that they secrete as well as to other possible growth stimulators. Furthermore, at sites where cell injury and necrosis occur, damaged smooth muscle cells could release FGF and in so doing could also stimulate neighbouring smooth muscle, the overlying endothelium or vascular channels within the lesions¹¹⁷. Their synthetic activity will determine the matrix content of the lesion, which in turn could interact at their surface and modify their capacity to respond to various agonists. Thus, the smooth muscle cell plays the principal role in the fibroproliferative component of this disease process.

Immune responses

In the atheromatous blood vessel, the presence of T lymphocytes, together with numerous macrophages, suggests that not only is

BOX 2 Macrophages in atherosclerosis

THE potential roles of the macrophage in atherogenesis. The reverse arrows between the T lymphocyte and macrophage suggest that some form of immune response may occur during atherogenesis. Interactions between T cells and macrophages can result in proliferation of each of these cell types through IL-2 and CSFs, respectively. All of the cells with which the macrophage can interact, namely the T lymphocyte, smooth muscle and endothelium, can present CSF to the macrophages to maintain cell viability and prevent apoptosis and cell death, and participate in further macrophage activation and replication. In addition, smooth muscle and endothelium can present antigens (Ag) at their surfaces and secrete chemoattractants for macrophages, including MCP-1 and oxLDL, as well as factors that can alter macrophage metabolism such as IL-1 or TNF α . When the macrophages are activated, they can produce an extraordinary number of biologically relevant molecules, some of which are listed in the box in relation to their capacity to induce or inhibit replication of endothelium, smooth muscle or macrophages themselves, as well as their capacity to make chemoattractants for each of these cell types. For colour coding see Fig. 1.



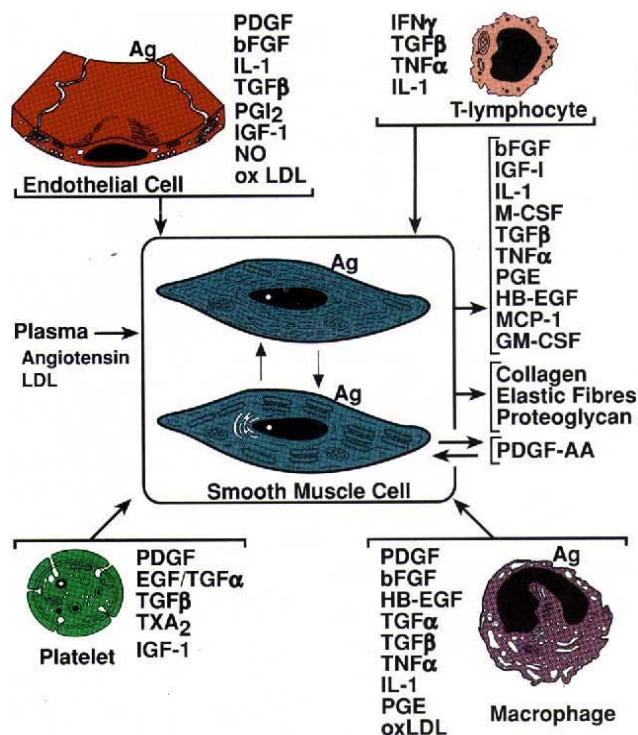
there an inflammatory reaction during the process of atherogenesis, but also that there may be an immunological reaction, and that this may be related to a specific antigen(s)^{1,3,24,28,102}. Although it has not been possible thus far to identify such an antigen, autoantibodies have been found to oxLDL. Furthermore, clonal expansion of lymphocytes does not appear to take place¹⁵. In addition, in the circumstances where there is indeed a known antigen, such as in the transplanted and rejected heart, the atherosclerotic lesions are different. They contain the same cellular elements as are found in plaques associated with diabetes, hypertension or hyperlipidaemia. But the plaques are concentric rather than eccentric in distribution and show a massive increase in lymphocyte and monocyte content. The high levels of HLA-DR present on the endothelium of the vessels from the transplanted hearts may account for the induction of the T-cell rejection response, but this does not explain the ubiquity of the T cell in most nontransplant-associated atherosclerotic lesions¹⁵.

Restenosis

The principal surgical approaches to the treatment of atherosclerosis are bypass grafting, endarterectomy, and percutaneous transluminal angioplasty (PTA). The failure rate due to restenosis after these approaches is extraordinarily high (30–50%), and it would seem that much of the restenosis is due to further inflammation, smooth muscle accumulation and thrombosis¹¹⁸. The principal animal model that has been used to study this process is balloon angioplasty of the normal rat carotid artery, studies of which have demonstrated that platelets, FGF, and PDGF are intimately involved in the intimal thickening that forms 7–21 days after injury^{56,119,120}. Antibodies against bFGF prevent the medial smooth muscle replication that occurs 24–48 hours after the medial cells are injured by the inflated balloon¹²⁰. Antibodies against platelets or PDGF induce a statistically significant diminution in the neointimal accumulation of smooth muscle cells that results 3–6 days after ballooning, and appear to block migration of cells from the media into the intima^{56,119}. Similarly, infusion of IFN γ ¹²¹ reduces the size of the neointima; thus, numerous factors may interact in this process. Factors responsible for the neointimal proliferation that occurs between days 7 and 21 are not yet known. Unfortunately, as useful as the rat model has been in studying these processes, it falls short of the process of restenosis postangioplasty in humans. This involves angioplasty of a pre-existing, occlusive lesion of atherosclerosis rather than a normal artery. The cells from the lesions of atherosclerosis undoubtedly respond differently in

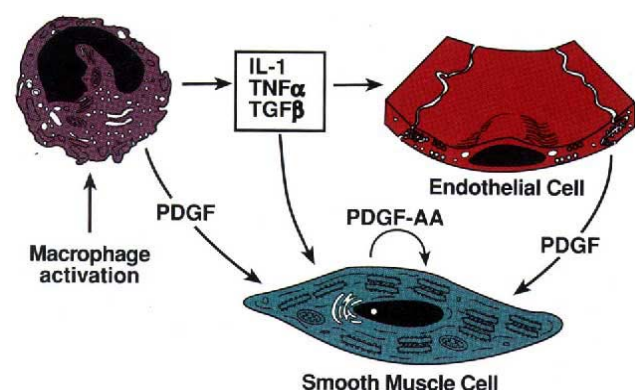
BOX 3 The two potentially different phenotypic states of the smooth muscle cell (centre)

In the synthetic phenotype it is presumed that smooth muscle cells can form connective tissue molecules as well as growth factors such as PDGF-AA, and can stimulate themselves as well as their neighbours. In their interactions with the overlying endothelium and neighbouring T lymphocytes, platelets and macrophages, smooth muscle cells can respond to the different cytokines, growth-regulatory molecules, and vasodilator and vasoconstrictor substances that can be generated from these cells, as well as substances from the plasma, such as angiotensin.



Thus, the genes that are expressed in the different phenotypic states by the smooth muscle (listed to the right), as well as those expressed by the neighbouring cells (listed next to each cell) in the artery wall that result from these cellular interactions will determine the outcome as to whether a lesion will progress or regress. For colour coding see Fig. 1.

BOX 4 The common mode of smooth muscle proliferation that can be induced through IL-1, TNF α or TGF β



EACH of these molecules can be formed by activated macrophages and, when they are exposed to endothelial or smooth muscle cells, each of these molecules will induce PDGF B- or A-chain gene expression, respectively. *In vitro* studies, the mitogenic activity induced in endothelium or smooth muscle by IL-1, TNF α or TGF β can be totally abolished by a neutralizing polyclonal antibody to PDGF. In addition to its capacity to form IL-1, TNF α and TGF β , the macrophage can directly form PDGF and thus directly induce smooth muscle cell replication as well. For colour coding see Fig. 1.

view of their different composition and history. Cells from advanced human lesions grow very poorly in culture and have growth characteristics of senescent cells¹²². Recent use of an atherectomy catheter has provided observations of segments of primary and restenotic human lesions, which support the notion that both smooth muscle proliferation and migration are important in the luminal narrowing that results during the restenotic process¹³. Analyses of such restenotic material by *in situ* hybridization have demonstrated increased expression of TGF β and PDGF in the restenosing lesions¹²³.

Future approaches

The ability to target appropriate molecules specifically to sites at increased risk represents an exciting approach to treatment both of the restenosis and of the primary atherosclerotic lesions. Antagonists to several of the growth-regulatory molecules, including PDGF A- and B-chains, TGF β , bFGF, and the cytokines, IL-1 and TNF α , represent valuable approaches towards treatment and prevention of restenosis and possibly of the atherosclerosis itself. Because cell migration and proliferation are critical events in both responses, the targeting of signalling pathways and elements that control them may yield exciting results. For example, antisense oligonucleotides to a series of proto-oncogenes expressed in cells, including *c-myc*, *c-fos* or *c-myc*¹²⁴, to cytosolic enzymes or other molecules involved in mitogenic signalling, and possibly to some of the contractile elements, such as particular forms of smooth muscle myosin, could potentially be used to control these events. It is already possible to increase the stability of antisense oligonucleotides and to apply or direct them to local sites where injury has occurred, permitting them selectively to affect processes associated with smooth muscle migration and proliferation¹²⁴.

Occlusive lesions of atherosclerosis in humans can be clinically reversed after aggressive treatment with lipid-lowering drugs^{125,126}. Thus, recognition of the importance of inflammation, smooth muscle and monocyte replication, and connective tissue formation, together with lipid accumulation, provides numerous entry points for selective interference and inhibition of the process of atherogenesis^{1,127}. LDL oxidation increases monocyte adherence, transmigration, and macrophage formation and activation. These events could in turn result in lipid accumulation, focal areas of necrosis covered by areas of smooth

muscle and macrophage proliferation, connective tissue formation, and all of the subcellular events that result in advanced lesions of atherosclerosis (Fig. 1). Thus, elucidation of each of these phenomena and development of specific means of their selective inhibition should provide numerous opportunities to diagnose, treat and prevent the process of atherosclerosis.

Maintenance of homeostasis and control of the local environment are critical factors involved in the prevention of disease. The aetiology and pathogenesis of atherosclerosis are tightly linked with the ubiquitous protective mechanisms associated with inflammation and repair. Should these become excessive and prolonged, they can both destroy sufficient tissue and induce such a chronic, inflammatory-fibroproliferative response that the intended protection becomes a disease entity in itself. Thus, optimization of the inflammatory response may represent a key to the development of new approaches to treatment and prevention of this disease process that is responsible for such enormous morbidity and mortality in developed countries. □

R. Ross is at the Department of Pathology, SM-30, University of Washington School of Medicine, Seattle, Washington 98195, USA.

- Ross, R. *N. Engl. J. Med.* **314**, 488-500 (1986).
- World Health Organization *WHO Tech. Rep. Ser.* **143**, 1-20 (1985).
- Stary, H. C. *Arteriosclerosis* **9** (suppl. 1), 119-132 (1989).
- McGill, H. C. Jr *Arteriosclerosis* **4**, 443-451 (1984).
- Faggiotto, A., Ross, R. & Harker, L. *Arteriosclerosis* **4**, 323-340 (1984).
- Faggiotto, A. & Ross, R. *Arteriosclerosis* **4**, 341-356 (1984).
- Masuda, J. & Ross, R. *Arteriosclerosis* **10**, 164-177 (1990).
- Masuda, J. & Ross, R. *Arteriosclerosis* **10**, 178-187 (1990).
- Rosenfeld, M. E., Tsukada, T., Gown, A. M. & Bondjers, G. *Arteriosclerosis* **7**, 9-23 (1987).
- Rosenfeld, M. E. et al. *Arteriosclerosis* **7**, 24-34 (1987).
- Gown, A. M., Tsukada, T. & Ross, R. *Am. J. Path.* **125**, 191-207 (1986).
- Jonasson, L., Holm, J., Skalli, O., Bondjers, G. & Hansson, G. K. *Arteriosclerosis* **6**, 131-138 (1986).
- Dartsch, P. C. et al. *Atherosclerosis* **80**, 149-157 (1989).
- Hansson, G. K., Seifert, P. S., Olsson, G. & Bondjers, G. *Arterioscl. Thromb.* **11**, 745-750 (1991).
- Libby, P. & Hansson, G. K. *Lab. Invest.* **64**, 5-15 (1991).
- Davies, M. J. & Thomas, A. N. *Engl. J. Med.* **310**, 1137-1140 (1984).
- Virchow, R. in *Gesammelte Abhandlungen zur Wissenschaftlichen Medizin* 458-463 (Meidinger Sohn, Berlin, 1856).
- French, J. E. *Int. Rev. exp. Pathol.* **5**, 253-353 (1966).
- Ross, R. & Glomset, J. A. *Science* **180**, 1332-1339 (1973).
- Ross, R. & Glomset, J. A. *N. Engl. J. Med.* **295**, 369-377, 420-425 (1976).
- Ross, R. & Harker, L. *Science* **193**, 1094-1100 (1976).
- Ross, R. *Arteriosclerosis* **1**, 293-311 (1981).
- Mora, R., Lupu, F. & Simionescu, N. *Atherosclerosis* **67**, 143-154 (1987).
- Munro, J. M. & Cotran, R. S. *Lab. Invest.* **58**, 249-261 (1988).
- Davies, P. F. *Lab. Invest.* **55**, 5-24 (1986).
- Gerrity, R. G., Goss, J. A. & Soby, L. *Arteriosclerosis* **5**, 55-66 (1985).
- Nilsson, J. *Atherosclerosis* **62**, 185-199 (1986).
- Thomas, W. A., Lee, K. T. & Kim, D. N. *Ann. N.Y. Acad. Sci.* **454**, 305-315 (1985).
- Tsukada, T., Rosenfeld, M. E., Ross, R. & Gown, A. M. *Arteriosclerosis* **6**, 601-613 (1986).
- Davies, M. J., Woolf, N., Rowles, P. M. & Pepper, J. *Br. Heart J.* **60**, 459-464 (1988).
- Ross, R., Raines, E. W. & Bowen-Pope, D. F. *Cell* **46**, 155-169 (1986).
- Raines, E. W., Bowen-Pope, D. F. & Ross, R. in *Handbook of Experimental Pharmacology: Peptide Growth Factors and their Receptors* (eds Sporn, M. B. & Roberts, A. B.) 173-262 (Springer, Berlin, 1990).
- Hart, C. E. et al. *Science* **240**, 1529-1531 (1988).
- Heldin, C.-H. & Westermark, B. *Cell Reg.* **1**, 555-566 (1990).
- Burgess, W. H. & Maciag, T. A. *Rev. Biochem.* **58**, 575-606 (1989).
- Klagsbrun, M. & Edelman, E. R. *Arteriosclerosis* **9**, 269-278 (1989).
- Folkman, J. et al. *Am. J. Path.* **130**, 393-400 (1988).
- Higashiyama, S., Abraham, J. A., Miller, J., Fiddes, J. C. & Klagsbrun, M. *Science* **251**, 936-939 (1991).
- Banskota, N. K., Taub, R., Zellner, K. & King, G. L. *Molec. Endocrin.* **3**, 1183-1190 (1989).
- Libby, P., Friedman, G. B. & Salomon, R. N. *Am. Rev. resp. Dis.* **140**, 1114-1117 (1989).
- Raines, E. W., Dower, S. K. & Ross, R. *Science* **243**, 393-396 (1989).
- Old, L. J. *Science* **230**, 630-632 (1985).
- Moses, H. L., Coffey, R. J. Jr, Leaf, E. B., Lyons, R. M. & Keski-Oja, J. *J. Cell Physiol.* (suppl. 5), 1-7 (1987).
- Sporn, M. B., Roberts, A. B., Wakefield, L. M. & de Crombrughe, B. *J. Cell Biol.* **105**, 1039-1045 (1987).
- Roberts, A. B. & Sporn, M. B. in *Handbook of Experimental Pharmacology: Peptide Growth Factors and their Receptors I* (eds Sporn, M. B. & Roberts, A. B.) 419-472 (Springer, Berlin, 1989).
- Moses, H. L., Yang, E. Y. & Pietenpol, J. A. *Cell* **63**, 245-247 (1990).
- McCaffrey, T. A. et al. *J. Cell Biol.* **109**, 441-448 (1989).
- Chaudhry, A. R. & D'Amore, P. A. *J. Cell Biochem.* **47**, 224-229 (1991).
- Battagay, E. J., Raines, E. W., Seifert, R. A., Bowen-Pope, D. F. & Ross, R. *Cell* **63**, 515-524 (1990).
- Stanley, E. R. *Meth. Enzym.* **116**, 564-587 (1985).
- Ylä-Herttuala, S. et al. *Proc. natn. Acad. Sci. U.S.A.* **88**, 5252-5256 (1991).
- Leonard, E. J. & Yoshimura, T. *Immun. Today* **11**, 97-101 (1990).
- Cathcart, M. K., Morel, D. W. & Chisolm, G. M. *J. Leuk. Biol.* **38**, 341-350 (1985).
- Rosenfeld, M. E., Palinski, W., Ylä-Herttuala, S. & Carew, T. E. *Toxicol. Path.* **18**, 560-571 (1990).
- Quinn, M. T., Parthasarathy, S. & Steinberg, D. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2805-2809 (1988).
- Ferns, G. A. et al. *Science* **253**, 1129-1132 (1991).
- Amento, E. P., Ehsani, N., Palmer, H. & Libby, P. *Arterioscl. Thromb.* **11**, 1223-1230 (1991).
- Steinbrecher, U. P., Zhang, H. F. & Loughheed, M. *Free Radical Biol. Med.* **9**, 155-168 (1990).
- Libby, P., Warner, S. J. C., Salomon, R. N. & Birnby, L. K. *N. Engl. J. Med.* **318**, 1493-1498 (1988).
- Hajjar, K. A., Hajjar, D. P., Silverstein, R. L. & Nachman, R. L. *J. exp. Med.* **166**, 235-245 (1987).

61. Wilcox, J. N., Smith, K. M., Williams, L. T., Schwartz, S. M. & Gordon, D. *J. clin. Invest.* **82**, 1134–1143 (1988).
62. Ross, R. *et al. Science* **248**, 1009–1012 (1990).
63. Barrett, T. B. & Benditt, E. P. *Proc. natn. Acad. Sci. U.S.A.* **85**, 2810–2814 (1988).
64. Rubin, K. *et al. Lancet* **i**, 1353–1356 (1988).
65. Raines, E. W. & Ross, R. *J. Cell Biol.* **116**, 533–543 (1992).
66. Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P. & Ornitz, D. M. *Cell* **64**, 841–848 (1991).
67. Rosenfeld, M. E. & Ross, R. *Arteriosclerosis* **10**, 680–687 (1990).
68. Gordon, D., Reidy, M. A., Benditt, E. P. & Schwartz, S. M. *Proc. natn. Acad. Sci. U.S.A.* **87**, 4600–4604 (1990).
69. Rosenfeld, M. E. *et al. Am. J. Path.* **140**, 291–300 (1992).
70. Shimano, H. *et al. J. Biol. Chem.* **265**, 12869–12875 (1990).
71. Gimbrone, M. A. Jr *Prog. Hemost. Thromb.* **3**, 1–28 (1976).
72. Steinberg, D. *Circulation* **84**, 1420–1425 (1991).
73. Springer, T. A. *Nature* **346**, 425–434 (1990).
74. Cybulsky, M. I. & Gimbrone, M. A. Jr *Science* **251**, 788–791 (1991).
75. Carlos, T. M. & Harlan, J. M. *Immun. Rev.* **114**, 5–28 (1990).
76. Ylä-Herttuala, S. *et al. J. clin. Invest.* **84**, 1086–1095 (1989).
77. Minor, R. L., Myers, P. R., Guerra, R. Jr, Bates, J. N. & Harrison, D. G. *J. clin. Invest.* **86**, 2109–2116 (1990).
78. Ylä-Herttuala, S. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 6959–6963 (1990).
79. McNally, A. K., Chisolm, G. M. III, Morel, D. W. & Cathcart, M. K. *J. Immun.* **145**, 254–259 (1990).
80. Furchtgott, R. F. *Circ. Res.* **53**, 557–573 (1983).
81. Lüscher, T. F. *Am. J. Hypertens.* **3**, 317–330 (1990).
82. Chester, A. H., O'Neill, G. S., Moncada, S., Tadjkarimi, S. & Yacoub, M. H. *Lancet* **336**, 897–900 (1990).
83. Botting, R. & Vane, J. R. *Arch. Mal. Coeur* **82**, 11–14 (1989).
84. Shimokawa, H. & Vanhoutte, P. M. *Circ. Res.* **64**, 900–914 (1989).
85. Dzau, V. J. *J. cardiovasc. Pharmac.* **15** (suppl. 5), S59–S64 (1990).
86. Owens, G. K. *Am. J. Physiol.* **257**, H1755–H1765 (1989).
87. Moncada, S. & Higgs, E. A. *Clin. Haematol.* **15**, 273–292 (1986).
88. Hajjar, D. P. *Am. J. Path.* **139**, 1195–1211 (1991).
89. Minick, C. R., Fabricant, C. G., Fabricant, J. & Litrenta, M. M. *Am. J. Path.* **96**, 673–700 (1979).
90. Rosenberg, R. D. *Fed. Proc.* **44**, 404–409 (1985).
91. Tolins, J. P., Shultz, P. J. & Raji, L. *Hypertension* **17**, 909–916 (1991).
92. Loskutoff, D. J. & Curriden, S. A. *Ann. N.Y. Acad. Sci.* **598**, 238–247 (1990).
93. Gordon, D., Augustine, A. J., Smith, K. M., Schwartz, S. M. & Wilcox, J. N. *Thromb. Haemostasis* **62**, 131 (1989).
94. Sawdey, M. S. & Loskutoff, D. J. *J. clin. Invest.* **88**, 1346–1353 (1991).
95. Goldstein, J. L., Ho, Y. K., Basu, S. K. & Brown, M. S. *Proc. natn. Acad. Sci. U.S.A.* **76**, 333–337 (1979).
96. Parthasarathy, S., Wieland, E. & Steinberg, D. *Proc. natn. Acad. Sci. U.S.A.* **86**, 1046–1050 (1989).
97. Rosenfeld, M. E. *et al. J. clin. Invest.* **87**, 90–99 (1990).
98. Carew, T. E., Schwenke, D. C. & Steinberg, D. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7725–7729 (1987).
99. Kita, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 5928–5931 (1987).
100. Shimokado, K. *et al. Cell* **43**, 277–286 (1985).
101. Libby, P. *et al. Am. J. Path.* **124**, 179–185 (1986).
102. Hansson, G. K., Jonasson, L., Seifert, P. S. & Stemme, S. *Arteriosclerosis* **9**, 567–578 (1989).
103. Schwartz, S. M., Heimark, R. L. & Majesky, M. W. *Physiol. Rev.* **70**, 1177–1209 (1990).
104. Hultgardh-Nilsson, A., Krondahl, U., Querol-Ferrer, V. & Ringertz, N. R. *Differentiation* **47**, 99–105 (1991).
105. Campbell, G. R., Campbell, J. H., Manderson, J. A., Horigan, S. & Rennick, R. E. *Arch. Path. Lab. Med.* **112**, 977–986 (1988).
106. Thyberg, J., Hedin, U., Sjölund, M., Palmberg, L. & Bottger, B. A. *Arteriosclerosis* **10**, 966–990 (1990).
107. Nagai, R., Kuro-o, M., Babji, P. & Periasamy, M. *J. Biol. Chem.* **264**, 9734–9737 (1989).
108. Yanagisawa, M. *et al. Nature* **332**, 411–415 (1988).
109. Majesky, M. W., Daemen, M. J. A. P. & Schwartz, S. M. *J. Biol. Chem.* **265**, 1082–1088 (1990).
110. Taubman, M. B. *et al. J. Biol. Chem.* **264**, 526–530 (1989).
111. Owen, N. E. *Am. J. Physiol.* **250**, C584–C588 (1986).
112. Morisaki, N., Kanzaki, T., Motoyama, N., Saito, Y. & Yoshida, S. *Atherosclerosis* **71**, 165–171 (1988).
113. Nilsson, J., von Euler, A. M. & Dalsgaard, C.-J. *Nature* **315**, 61–63 (1985).
114. Palmberg, L., Claesson, H.-E. & Thyberg, J. *J. Cell Sci.* **88**, 151–159 (1987).
115. Moncada, S. & Higgs, E. A. (eds) *Nitric Oxide from L-arginine: A Bioregulatory System* (Excerpta Medica, Amsterdam, 1990).
116. Sjölund, M., Hedin, U., Seijersen, T., Heldin, C.-H. & Thyberg, J. *J. Cell Biol.* **106**, 403–413 (1988).
117. Baird, A. & Böhlen, P. in *Peptide Growth Factors and their Receptors I* (eds Sporn, M. B. & Roberts, A. B.) 369–418 (Springer, Berlin, 1990).
118. Ferrell, M., Fuster, V., Gold, H. K. & Chesebro, J. H. *Circulation* **85**, 1630–1631 (1992).
119. Fingler, J., Johnson, R., Clowes, A. W., Majesky, M. W. & Reidy, M. A. *Proc. natn. Acad. Sci. U.S.A.* **86**, 8412–8416 (1989).
120. Lindner, V. & Reidy, M. A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 3739–3743 (1991).
121. Hansson, G. K. & Holm, J. *Circulation* **84**, 1266–1272 (1991).
122. Ross, R., Wight, T. N., Strandness, E. & Thiele, B. *Am. J. Path.* **114**, 79–93 (1984).
123. Nikol, S. *et al. J. clin. Invest.* **90**, 1582–1592 (1992).
124. Simons, M., Edelman, E. R., DeKeyser, J.-L., Langer, R. & Rosenberg, R. D. *Nature* **359**, 67–70 (1992).
125. Brown, B. G. *et al. N. Engl. J. Med.* **323**, 1289–1298 (1990).
126. Kane, J. P., Malloy, M. J., Ports, T. A., Phillips, N. R., Diehl, J. C. & Havel, R. H. *J. Am. med. Assoc.* **264**, 3007–3012 (1990).
127. Fuster, V., Badimon, L., Badimon, J. J. & Chesebro, J. H. *Engl. J. Med.* **326**, 242–250 (1992).

ACKNOWLEDGEMENTS. I thank E. W. Raines and M. E. Rosenfeld, who, together with many postdoctoral fellows, played a critical role in much of the research and made many suggestions. N. Woolf and M. J. Davies for the scanning electron micrographs of the human coronary arteries and J. Masuda and M. Sasahara for contributions from their studies of *Macaca nemestrina*. I am particularly indebted to K. Carroll and M. Bohidar for the diagrams in Fig. 1 and Boxes 1–4 and to B. Droker for preparation of the manuscript. This review was supported in part by a grant from the National Heart, Lung, and Blood Institute, NIH, and an unrestricted grant for cardiovascular research from Bristol-Myers Squibb Company.

SUPPLEMENTARY INFORMATION. A full bibliography was submitted with the manuscript and is available on request to the author or from the London editorial office of *Nature*.

Oxygen isotope evidence for recycled crust in the source of EM-type ocean island basalts

Jon D. Woodhead*, Peter Greenwood†, Russell S. Harmon† & Peter Stoffers‡

* Research School of Earth Sciences, Australian National University, Canberra, Australian Capital Territory 0200, Australia

† NERC Isotope Geosciences Laboratory, Kingsley Dunham Centre, Keyworth, Nottingham NG12 5GG, UK

‡ Geologische Institut der Universität Kiel, 40–60 Olshausenstrasse, 2300 Kiel, Germany

The oxygen isotope compositions of fresh volcanic glasses from submarine volcanoes of 'enriched mantle' (EM-I and EM-II) type appear to require the presence of subducted crustal components in the source of these lavas. These data confirm earlier inferences from trace elements and radiogenic isotopes, and also provide a rare opportunity to estimate the proportion of subducted material that reaches the source region of ocean island basalts.

ALTHOUGH recycled crustal components have been identified in the source of convergent margin magmas, their fate beyond Benioff zone depths is still a matter of conjecture. Since Hofmann and White¹ first proposed that subducted oceanic crust, with or without sediment, could contribute to the mantle source of plume-derived ocean island basalts (OIB), the search for the geochemical signature of such components in basaltic rocks has been dominated by radiogenic isotope and trace element studies. The compositions of most OIB seem to be explicable in terms of multi-component mixing between up to five distinct end-members, four of which have been defined from extrapolation of existing data arrays in isotope space, as shown

in Fig. 1 (the existence of the fifth, primitive mantle, is largely conjectural)². These have been termed 'MORB', a component similar to the source of mid-ocean-ridge basalt, 'HIMU', an acronym for high U/Pb ratio, characterized by very radiogenic Pb isotope compositions, and two 'enriched mantle' components, EM-I and EM-II³. Of these, the EM components possess radiogenic isotope compositions suggesting the influence of crustal material in their sources, having apparently experienced long-term evolution under conditions of high Rb/Sr, and low Sm/Nd ratio. But these interpretations are non-unique, and alternative models have been proposed to account for the Sr, Nd and Pb-isotope compositions and patterns

of incompatible-element enrichment seen in EM-type reservoirs. For example, it has been suggested that subcontinental lithospheric mantle is involved^{4,5}.

Stable-isotope tracers, especially oxygen, have been largely ignored in this context because of the uncertainties associated with surficial alteration processes, as most volcanic islands are old (>0.5–1 Myr), and are often areas of high rainfall, and thus may have had their bulk O-isotope composition modified by hydration and/or weathering. In addition, the oceanic upper-mantle reservoir, as defined by analyses of MORB⁶, shows extreme uniformity at $\delta^{18}\text{O} \approx +5.7 \pm 0.2\text{‰}$ (all data are quoted relative to the SMOW standard). However, dredging by the German research vessel *Sonne* has provided fresh, recently erupted, volcanic glasses, free from obvious effects of alteration, from sites of active volcanism on the Pitcairn (EM-I)⁷ and Society (EM-II)⁸ hotspots in Polynesia, both of which have been suggested to have geochemical characteristics influenced by recycled crustal components^{9,10}. This has prompted us to re-appraise the use of oxygen isotopes in the study of crustal recycling. These glasses, which form the chilled margins to volcanic pillows and sheet flows, have oxygen isotope compositions that range from 5.8 to 7.4‰, clearly distinct from MORB. Detailed consideration of these data seems to require up to 9% recycled crustal oxygen in the mantle source of these OIB.

Oxygen isotope constraints

There are significant $\delta^{18}\text{O}$ variations in both sample suites (Table 1), extending to values substantially higher than the MORB

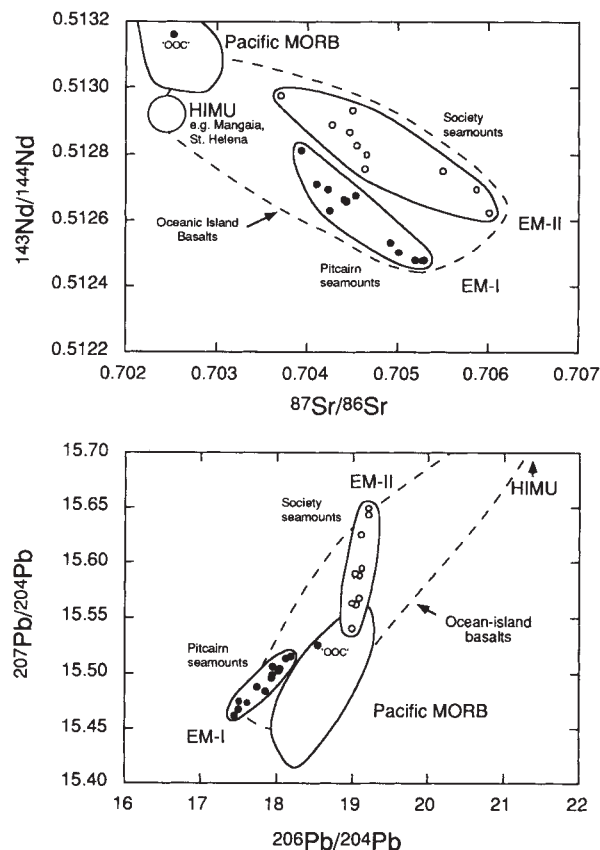


FIG. 1 Radiogenic isotope plots illustrating the geochemical characteristics of MORB, HIMU and EM-type mantle sources, as defined from extrapolation of existing data arrays in isotope space (the ocean-island basalt field encompasses most known analyses). Also shown are data for the Pitcairn and Society seamount glasses^{9,10} used in this study as representative of EM-I and EM-II sources respectively. The old oceanic crust (OOC) datum represents an analysis of a low-potassium tholeiite derived from an old seamount in the Pitcairn region which is interpreted to have formed ~20–25 Myr ago near a mid-ocean ridge⁹.

average of +5.7‰, which is generally considered as the norm for the oceanic mantle. Samples from the Pitcairn region form very pronounced curvilinear arrays in plots of radiogenic against oxygen isotopes (Fig. 2), consistent with a simple, two-component source mixing process (typical mixing hyperbolae are plotted in Fig. 2 for comparison). In this case, one end-member tends towards MORB-like compositions with $\delta^{18}\text{O} \approx +5.7\text{‰}$, $^{143}\text{Nd}/^{144}\text{Nd} > 0.5128$, $^{87}\text{Sr}/^{86}\text{Sr} < 0.704$ and $^{206}\text{Pb}/^{204}\text{Pb} > 18.25$, whereas the other, interpreted as the 'plume component', has $\delta^{18}\text{O} > +7.4\text{‰}$, $^{143}\text{Nd}/^{144}\text{Nd} < 0.51245$, $^{87}\text{Sr}/^{86}\text{Sr} > 0.7055$ and $^{206}\text{Pb}/^{204}\text{Pb} < 17.5$. Further, the curvature of the arrays indicates that the latter component has a higher Nd/O ratio. Within the Society data set, in contrast, no pronounced covariation occurs between radiogenic and oxygen isotopes

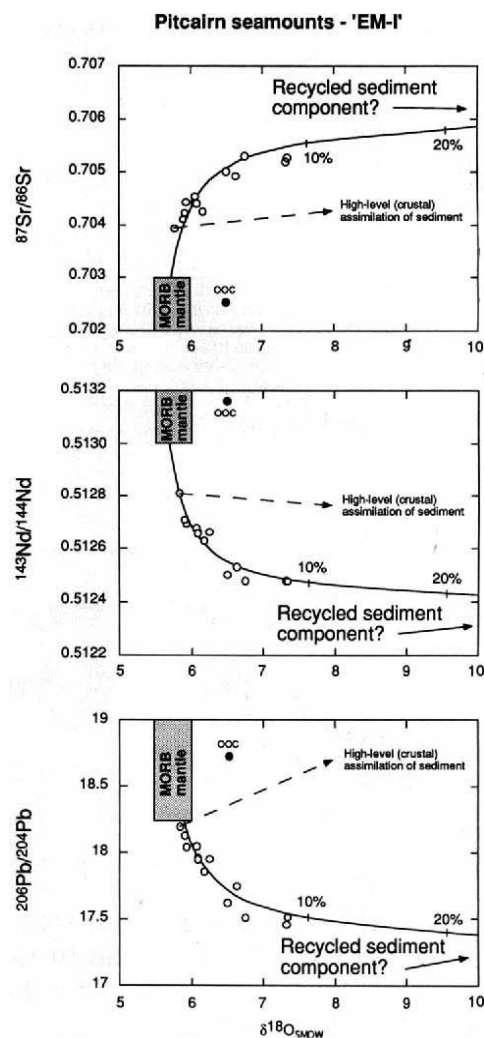


FIG. 2 $^{87}\text{Sr}/^{86}\text{Sr}$, $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{206}\text{Pb}/^{204}\text{Pb}$ against $\delta^{18}\text{O}$ for 'EM-I' type glasses from the Pitcairn seamounts. All the data extend to $\delta^{18}\text{O}$ higher than the MORB average of $5.7 \pm 0.2\text{‰}$, and form pronounced correlations with radiogenic isotopes. Mixing hyperbolae are shown between MORB-source mantle and a recycled sediment component with respectively, $\delta^{18}\text{O}$ of +5.7‰, $^{87}\text{Sr}/^{86}\text{Sr}$ of 0.7027 and 0.706, $^{143}\text{Nd}/^{144}\text{Nd}$ of 0.5130 and 0.5124, and $^{206}\text{Pb}/^{204}\text{Pb}$ of 18.8 and 17.3. Concentration parameters (concentration of element in mantle: concentration in sediment) were chosen to provide the closest approximation to the data array: Sr: 1:50, Nd: 1:70, Pb: 1:60. Also shown are typical vectors which might be expected for assimilation of modern sedimentary material within the crust by an ascending OIB magma; such a model is clearly inappropriate to the data. In our preferred interpretation the data are consistent with the incorporation of up to 9% crustal oxygen into the mantle source of these magmas (datum marks show percentage of crustal oxygen).

(Fig. 3) but once again all samples have $\delta^{18}\text{O} > 6.0\%$. No correlations are noted in either suite between oxygen and helium isotope data (noble gas analyses for the Pitcairn and Society glasses: M. Honda, personal communication, and ref. 11, respectively).

Although we would contend that these O-isotope variations are inherited from the mantle source of these lavas, it is important first to establish whether the observed values can be related to secondary, post-eruptive processes or magmatic phenomena occurring in high-level magma chambers.

Secondary alteration processes. It is well known that hydrothermal alteration can substantially modify oxygen isotope ratios and that the outcome of such a process is strongly temperature-dependent (see, for example, ref. 12). Low-temperature alteration, resulting from hydration and/or submarine weathering, affects highly polymerized, compositionally evolved lavas more than their mafic counterparts and produces strong ^{18}O enrichment, as a result of either a fractionation of +20 to 25‰ during hydration (favouring the heavy isotope) or addition of a smectite component of similarly high $\delta^{18}\text{O}$. Alternatively, direct O-isotope diffusion exchange may take place as a result of fluid-

rock interaction: if this occurred while the rocks were still hot ($>350^\circ\text{C}$), the result would be to pull the rock $\delta^{18}\text{O}$ values down towards the seawater value of 0‰, whereas low-temperature water-rock interaction, which is essentially equivalent to hydration, would lead to strong ^{18}O enrichment, particularly in silicic glasses.

Obvious secondary products such as amphibole, epidote and smectites were absent from the samples studied and, by careful hand-picking, we believe that we have avoided any non-pristine material. In addition, there is no geochemical evidence for sea water interaction in these samples. Whatever form of interaction one might envisage, this would be expected to introduce radiogenic Sr, coupled with O-isotope modification, so we should see correlated Sr-O isotopic variations but decoupled Sr-Nd and Nd-O systematics, as the introduction of sea water will not modify the Nd-isotope composition of the magma. The marked correlation between Nd and oxygen isotopes observed within the Pitcairn data (Fig. 2) therefore argue against seawater interaction (note also that strong leaching of these glasses before analysis^{9,10} has no effect on radiogenic isotope compositions,

TABLE 1 Oxygen isotope and other compositional data for EM-type OIB glasses

<i>Pitcairn seamounts-EM-I suite</i>								
Sample	SiO ₂	MgO	⁸⁷ Sr/ ⁸⁶ Sr	¹⁴³ Nd/ ¹⁴⁴ Nd	²⁰⁶ Pb/ ²⁰⁴ Pb	²⁰⁷ Pb/ ²⁰⁴ Pb	²⁰⁸ Pb/ ²⁰⁴ Pb	$\delta^{18}\text{O}$
Volcano 2								
33DS2	49.63	5.70	0.704107	0.512708	18.124	15.513	38.936	5.9
45DS1	53.50	3.39	0.704920	0.512530	17.743	15.487	38.966	6.6
46DS2	48.00	7.37	0.704230	0.512690	18.037	15.502	38.939	5.9
48DS6	49.01	5.74	0.704251	0.512629	17.853	15.484	38.611	6.2
49DS1	46.73	14.80	0.705010	0.512501	17.619	15.473	38.859	6.5
51DS1	49.77	9.14	0.704431	0.512661	17.948	15.499	38.809	6.0
51DS2	50.01	9.50	0.704431	0.512657	17.944	15.496	38.805	6.1
51DS4	54.21	2.76	0.703932	0.512811	18.191	15.515	38.898	5.8
51DS7	50.41	9.05	0.704407	0.512663	17.956	15.506	38.826	6.1
52DS1	54.15	1.70	0.705193	0.512478	17.454	15.461	38.993	7.3
55DS6	48.81	6.49	0.704539	0.512676	18.041	15.504	38.940	6.1
Volcano 5								
57DS1	54.78	2.56	0.705267	0.512476	17.510	15.474	38.890	7.4
57DS6	60.55	0.99	0.705296	0.512478	17.507	15.467	38.876	6.8
<i>Society seamounts-EM-II suite</i>								
Sample	SiO ₂	MgO	⁸⁷ Sr/ ⁸⁶ Sr	¹⁴³ Nd/ ¹⁴⁴ Nd	²⁰⁶ Pb/ ²⁰⁶ Pb	²⁰⁷ Pb/ ²⁰⁴ Pb	²⁰⁸ Pb/ ²⁰⁴ Pb	$\delta^{18}\text{O}$
Mehetia								
P1-1	45.30	7.80	0.704655	0.512799	19.095	15.567	38.949	6.3
P3-4	43.06	5.60	0.704276	0.512889	19.057	15.561	38.751	6.1
Teahitia								
D2-1	44.05	12.14	0.704466	0.512866	19.098	15.587	38.876	6.4
3-1	45.35	9.28	0.704640	0.512755	19.128	15.594	38.890	6.0
3-3	44.15	12.33	0.704547	0.512827	19.037	15.589	38.809	6.2
9-1	47.15	9.92	0.705506	0.512749	19.117	15.624	38.983	6.0
81-3	59.70	1.08	0.704504	0.512930	18.993	15.563	38.681	6.4
Rocard								
4-2	58.05	1.30	0.706017	0.512622	19.212	15.643	39.107	6.4
4-3	57.73	1.32	0.705870	0.512693	19.221	15.649	39.136	6.7
Moua Pihaa								
29-1	42.74	5.45	0.703709	0.512974	18.993	15.540	38.744	6.1
Statistics for duplicate analyses (samples and NBS28 standard)								
33DS2	n=2	$\bar{x}=5.91 \pm 0.05 (1\sigma)$						
9-1	n=2	$\bar{x}=6.04 \pm 0.08 (1\sigma)$						
51DS2	n=5	$\bar{x}=6.07 \pm 0.05 (1\sigma)$						
NBS28		$\bar{x}=9.62 \pm 0.15 (1\sigma)$						

Hand-picked glass chips were ground to a fine sand in an agate pestle and mortar. Stable isotope analyses were done at the NERC Isotope Geosciences Laboratory, UK. Oxygen was liberated from weighted aliquots by reaction with BrF_5 and converted to CO_2 by passing over a platinized graphite rod, heated by radiofrequency induction before mass spectrometric analysis. Results are quoted in the $\delta^{18}\text{O}$ notation, as per mil deviations from the SMOW standard. Duplicate analyses of samples show a mean variance of ± 0.06 . Radiogenic isotope analyses were done at the Research School of Earth Sciences, Canberra (Pitcairn seamounts) and Centre de Recherches Pétrographique et Géochimique, Nancy (Society seamounts). SiO₂ and MgO values are based on X-ray fluorescence data for co-existing (crystalline) pillow interiors from which the glassy rims were removed for isotope analysis. All data, except oxygen isotopes, reproduced from refs 9 and 10: analytical details can be found therein.

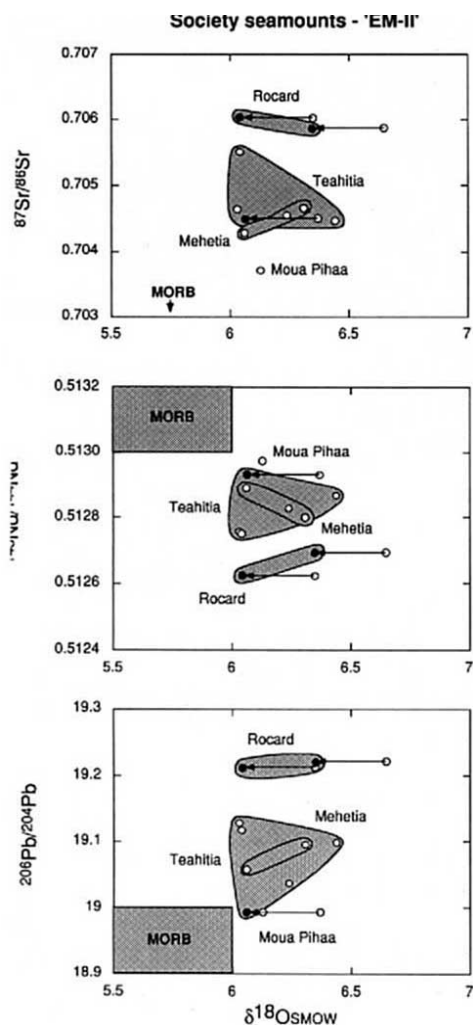


FIG. 3 $^{87}\text{Sr}/^{86}\text{Sr}$, $^{143}\text{Nd}/^{144}\text{Nd}$ and $^{206}\text{Pb}/^{204}\text{Pb}$ against $\delta^{18}\text{O}$ for 'EM-II' type glasses from the Society seamounts. Corrections for crystal fractionation effects have been applied to three samples which are relatively siliceous: these are shown by an arrow joining corrected (●) and uncorrected (○) data. In contrast to the Pitcairn glasses, no distinct correlations occur with radiogenic isotopes in this case but, once again, the $\delta^{18}\text{O}$ values are generally high, in comparison to MORB.

suggesting that these, and by inference the oxygen data, are primary). Finally, direct seawater interaction is likely to be diffusion-controlled and is therefore not considered an effective means of adding water to a magma¹³.

High-level magmatic processes. Simple crystal fractionation can result in an increase in $\delta^{18}\text{O}$ from primitive basaltic materials of ~0.3–0.4‰ per 10 wt% silica increase (see for example refs 14, 15). As some of the samples used in this study are non-basaltic, oxygen isotope variations are plotted against silica as an index of differentiation in Fig. 4. No systematic variations are observed within the Pitcairn suite, but there is a suggestion of an increase in $\delta^{18}\text{O}$ in the most evolved Society samples. For these three rocks, we therefore applied a small correction to offset the effects of fractionation (see Fig. 3). In view of the general lack of correlation in Fig. 4, and the very considerable degrees of fractionation required to attain the elevated oxygen isotope ratios observed in these suites, it appears unlikely that these can be related simply to crystal fractionation processes.

Assimilation of hydrothermally altered oceanic crust or sediment during magma transit or storage could produce ^{18}O enrichment in derivative magmas. The correlations we see are unlikely to result from assimilation of altered oceanic crust, as $\delta^{18}\text{O}$ and

Nd would not be expected to covary. James¹⁶ argues that the shape of mixing hyperbolae in plots such as those of Fig. 2 helps to distinguish between contamination at source (subducted, recycled sediment) and assimilation by a magma ascending through oceanic crust and sediment. These distinctions arise from the contrasting concentrations of Sr, Nd and Pb in crust, mantle and mantle-derived melt, respectively, which in turn control the shape of the mixing hyperbolae. The degree and sense of curvature of the hyperbolae in Fig. 2 indicate a source contamination process is most appropriate for these data (the dashed line in this figure would be more typical of a high-level (shallow) assimilation process where elemental concentrations are similar in both end-members). Finally, the isotopically enriched, high- $\delta^{18}\text{O}$ component seen in this data set must also have very unradiogenic Pb-isotope ratios with $^{206}\text{Pb}/^{204}\text{Pb} < 17.5$, unlike any known present-day sediment or indeed oceanic crust (for example the point labelled OOC in Fig. 2). Such a composition is, however, consistent with the involvement of ancient recycled sediment^{17–20}. These considerations suggest to us that the elevated oxygen isotope ratios observed in at least the Pitcairn suite, and possibly the Society data also, are primary and do not arise from water-magma interaction or late-stage alteration.

Crustal recycling and mass budgets

It is unlikely that the correlated variations in radiogenic and oxygen isotopes observed in these OIB suites are the result of intra-mantle differentiation processes, as the range of oxygen isotope variation observed is well outside that predicted on the basis of different temperatures or extents of partial melting under equilibrium conditions (oxygen isotope fractionation factors are small at mantle temperatures)²¹. The only sites of major oxygen isotope fractionation on the Earth today are within the crust-hydrosphere-atmosphere system. Thus, the data indicate that the high $\delta^{18}\text{O}$ values in these mantle-derived magmas result from the introduction of ^{18}O -enriched crustal material into the mantle source of EM-type OIB, which we would suggest then mix with a MORB-like reservoir ($\delta^{18}\text{O} \approx +5.7\text{‰}$), either during ascent or on contact with the oceanic lithosphere. One can envisage several different ways of introducing this crustal signature into the mantle: remobilization of metasomatized mantle wedge material from above subduction zones²², delamination of the sub-continental lithosphere⁴ (the formation of which may also ultimately be related to subduction zone processes), or direct addition of subducted oceanic sediments to the mantle source¹⁷. As the trace element geochemistry of OIB is so unlike that of convergent margin magmas, and the $\delta^{18}\text{O}$ of the assumed plume component is higher than that observed in typical primitive arc rocks (see, for example, ref. 15), we would suggest that the first two possibilities are less likely and that subduction of a small quantity of crustal material with $\delta^{18}\text{O} > 15\text{‰}$ into the mantle source most readily explains the data.

Oxygen isotopes provide robust and unique constraints on the proportion of any such material involved. In radiogenic systems, in addition to uncertainties in the isotopic composition of mixing end-members, there frequently is little constraint on the elemental concentrations in one or more of these end-members, thus making precise estimation of mixing proportions impossible. But as oxygen constitutes ~50% of both the Earth's mantle and its continental crust, and varies in different rock types only within narrow limits, estimation of mixing proportions in this case is only sensitive to the assumed isotopic composition of the end-members. It is also reasonable to assume that the isotopically depleted component (lithosphere or asthenosphere) has $\delta^{18}\text{O} \approx +5.7\text{‰}$, similar to MORB or MORB-source mantle. For ^{18}O -rich pelagic sediments, those most likely to be subducted into the deep mantle, $\delta^{18}\text{O}$ values will be high (+15 to +30‰, and averaging +25‰ for nine Pacific sediment analyses from refs 15 and 23). Thus, if we now assume a maximum $\delta^{18}\text{O}$ in the subducted component of +25‰,

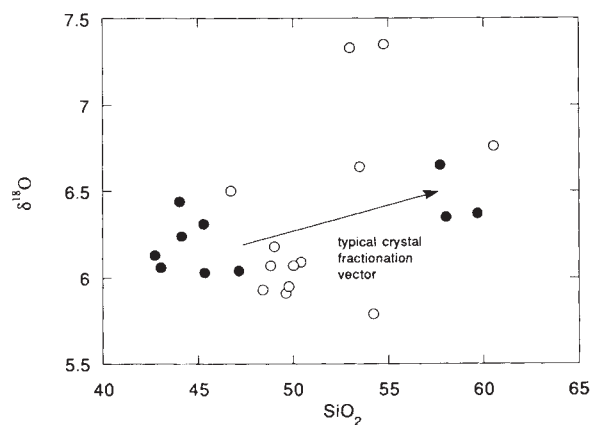


FIG. 4 $\delta^{18}\text{O}$ against silica for glasses from the Pitcairn seamounts, southeast Pacific (○) and Society seamounts, south central Pacific (●). Also shown is a typical crystal fractionation vector (0.3‰ per 10 wt% silica increase) from ref. 14.

equivalent to subducting pure pelagic sediment with no input from the downgoing oceanic crust, this provides a minimum estimate for the proportion of crustal oxygen involved in the OIB source (see datum marks on mixing hyperbolae in Fig. 2). For the most plume-dominated samples in this study (those with highest $^{87}\text{Sr}/^{86}\text{Sr}$ or lowest $^{143}\text{Nd}/^{144}\text{Nd}$), this corresponds to ~9% crustal oxygen contribution. This may itself be a mixture of heavy oxygen from both subducted sediments and hydrothermally altered oceanic crust. The radiogenic Nd isotope compositions suggest a substantial proportion of the former because, as noted above, seawater alteration has a limited effect on Nd isotope ratios.

Implications for the origin of OIB

It is clear from mass balance calculations²⁴, relating to the composition of the subducting slab and its contribution to convergent margin magma genesis, that considerable quantities of oceanic crust and sediment must have passed beyond Benioff zone depths into the upper mantle over the subduction history of the Earth. Mathematical models of convective stirring processes²⁵ suggest that geochemical heterogeneities, such as those generated by subduction, can survive within the mantle for billions of years, possibly with their O-isotope signatures intact²⁶. Thus, it is not unreasonable to invoke up to 9% recycled crustal oxygen in the OIB mantle source volume (as distinct from the whole mantle). This is, however, a result of great significance, as it implies that the patterns of incompatible element enrichment commonly observed in these OIB mainly result from incorporation of recycled material into their mantle source. Many authors have suggested that the trace element

compositions of EM-type OIB are consistent with oceanic crust and sediment involvement^{9,10,17,20}, but our results appear to require such a contribution.

The fact that good correlations with radiogenic tracers are observed in the Pitcairn suite and not within the Society data is clearly also of importance. The preservation of such correlations within the Pitcairn lavas suggests that mixing (but not necessarily recycling) was a fairly recent process, otherwise these arrays would have been destroyed or at least severely degraded by radiogenic isotope growth over the age of the source (Pb isotope modelling^{19,20,27} indicates that only recycled pelagic sediments older than at least 1.5 Gyr will result in compositions appropriate to the EM-I component, as the low U/Pb and high Th/U ratios can only produce a significant effect if considerable time has elapsed). This is consistent with mixing during diapir ascent, or interaction of magmas with the lithosphere before eruption. Within the Society data set (excluding Moua Pihaa¹⁰), correlations are observed between Sr, Nd and Pb in a manner similar to Pitcairn⁹, and thus the lack of covariation of these isotopes with the oxygen data may be indicating a much more ancient mixing event, possibly in the source.

In previous O-isotope determinations²⁸ for oceanic islands (whole rocks and mineral separates), the few samples from EM-type OIB show $\delta^{18}\text{O}$ generally higher than MORB and are thus consistent with the data presented above. We conclude that, in contrast to the low $\delta^{18}\text{O}$ values (+4.8–5.5‰) found in tholeiitic glasses from Hawaii²⁹, most fresh volcanic glasses from EM-type oceanic islands have relatively high $\delta^{18}\text{O}$ requiring small proportions (up to 9%) of recycled crustal oxygen in their mantle source. □

Received 4 November 1992; accepted 6 April 1993.

- Hofmann, A. W. & White, W. M. *Earth planet. Sci. Lett.* **57**, 421–436 (1982).
- White, W. M. *Geology* **13**, 115–118 (1985).
- Zindler, A. & Hart, S. R. *Rev. Earth planet. Sci.* **14**, 493–571 (1986).
- McKenzie, D. & O'Nions, R. K. *Nature* **301**, 229–231 (1985).
- Hawkesworth, C. J., Mantovani, M. S. M., Taylor, P. N. & Palacz, Z. *Nature* **322**, 356–359 (1986).
- Ito, E., White, W. M. & Göpel, C. *Chem. Geol.* **62**, 157–176 (1987).
- Scientific Party of cruise SO-65 of S.D. *Sonne Mar. Geol.* **95**, 51–55 (1990).
- Chiminee, J. I. *et al. Mar. geophys. Res.* **11**, 27–50 (1989).
- Woodhead, J. D. & Devey, C. *Earth planet. Sci. Lett.* (in the press).
- Devey, C. W. *et al. J. geophys. Res.* **95**, 5049–5066 (1990).
- Staudacher, T. & Allegre, C. J. *Earth planet. Sci. Lett.* **93**, 210–222 (1989).
- Muehlenbachs, K. & Clayton, R. N. *J. geophys. Res.* **81**, 4365–4369 (1976).
- Shaw, H. in: Hofmann, A. W., Giletti, B. J., Yoder, H. S. Jr & Jund, R. N. (eds) *Geochemical Transport and Kinetics*, 139–170 (Carnegie Inst. Washington Publ. No. 635, 1974).
- Sheppard, S. M. F. & Harris, C. *Contrib. Mineral. Petrol.* **91**, 74–81 (1985).
- Woodhead, J. D., Harmon, R. S. & Fraser, D. G. *Earth planet. Sci. Lett.* **83**, 39–52 (1987).
- James, D. E. A. *Rev. Earth planet. Sci.* **9**, 311–344 (1981).

- Weaver, B. L. *Earth planet. Sci. Lett.* **104**, 381–397 (1991).
- Ben Othman, D., White, W. M. & Patchett, J. *Earth planet. Sci. Lett.* **94**, 1–21 (1989).
- Woodhead, J. D. & McCulloch, M. T. *Earth planet. Sci. Lett.* **94**, 257–273 (1989).
- Chauvel, C., Hofmann, A. W. & Vidal, P. *Earth planet. Sci. Lett.* **110**, 99–119 (1992).
- Ionov, D. A., Harmon, R. S., France-Lanord, C. & Aschepkov, I. V. *Geochim. cosmochim. Acta* (in the press).
- Sun, S.-S. & McDonough, W. F. in: *Magmaism in the Ocean Basins* (eds Saunders, A. D. & Norry, M. J.) *Geol. Soc. Spec. Publ.* **42**, 313–345 (1989).
- Ito, E. & Stern, R. J. *Earth planet. Sci. Lett.* **76**, 312–320 (1985).
- White, W. M. & Patchett, J. *Earth planet. Sci. Lett.* **67**, 167–185 (1984).
- Davies, G. F. *Earth planet. Sci. Lett.* **99**, 94–109 (1990).
- Gregory, R. T. & Taylor, H. P. *Contrib. Mineral. Petrol.* **93**, 124–135 (1986).
- Weaver, B. L., Wood, D. A., Tarney, J. & Joron, J.-L. *Geology* **14**, 275–278 (1986).
- Sheppard, S. M. F. *Rev. Mineral.* **16**, 319–372 (1986).
- Garcia, M. O., Muenow, D. W. & Aggrey, K. E. *J. geophys. Res.* **94**, 10525–10538 (1989).

ACKNOWLEDGEMENTS. Research cruises in Polynesia were funded by German BMFT grants to P.S., and J.W. gratefully acknowledges funding from an Australian Research Council QEII Fellowship. We thank M. McCulloch and S. Eggins for enthusiastic discussion and Bill White for constructive comments.

Interfacial activation of the lipase–procolipase complex by mixed micelles revealed by X-ray crystallography

Herman van Tilbeurgh*, Marie-Pierre Egloff*, Chrislaine Martinez*, Nathalie Rugani†, Robert Verger‡ & Christian Cambillau*

* LCCMB-CNRS, Faculté de Médecine Nord, Boulevard P. Dramard, 13326 Marseilles Cedex 15, France

† Institut de Chimie Biologique, Faculté des Sciences Saint Charles, 3 Place Victor Hugo, 13331 Marseilles, France

‡ Laboratoire de Lipolyse Enzymatique, GDR1000 du CNRS, 31 Chemin Joseph Aiguier, 13402 Marseilles Cedex 9, France

The three-dimensional structure of the lipase–procolipase complex, co-crystallized with mixed micelles of phosphatidylcholine and bile salt, has been determined at 3 Å resolution by X-ray crystallography. The lid, a surface helix covering the catalytic triad of lipase, adopts a totally different conformation which allows phospholipid to bind to the enzyme's active site. The open lid is an essential component of the active site and interacts with procolipase. Together they form the lipid–water interface binding site. This reorganization of the lid structure provokes a second drastic conformational change in an active site loop, which in its turn creates the oxyanion hole (induced fit).

OUR understanding of the action of lipases has made much progress as a result of the resolution of the three-dimensional structure of three lipases, one of mammalian and two of microbial origin^{1–3}. Although these lipases share no sequence homology, they have in common a core consisting of a well conserved β -sheet, also present in the structure of esterases and alkane dehalogenase (the α/β -hydrolase fold)⁴. These lipase structures have confirmed the existence of a triad at the active site, made up of the amino-acid residues of Ser, His and Asp or Glu, like the catalytic triads found in the serine proteases^{1–3}. The importance of these three amino acids has been confirmed in pancreatic, hepatic and lipoprotein lipases by site-directed mutagenesis^{5–8}.

The three known lipase structures have their active site hidden from the bulk solvent by a lid, or flap. In the case of *Rhizomucor miehei* lipase and pancreatic lipase, the lid consists of an amphipathic helix². In lipase from *Geotrichum candidum*, probably two surface helices have to move to give full access to the active-site serine³. The structure of *Rhizomucor* lipase with a covalently bound inhibitor indicates the nature of the conformational change induced upon binding to the substrate: the helical lid rolls back upon the body of the molecule, augmenting substantially the hydrophobicity around the active site^{9,10}.

It was postulated that this lid movement is intimately related to the phenomenon of interfacial activation, namely the increase in lipase activity in the presence of a lipid–water interface as compared with a monomeric substrate solution. In the case of pancreatic lipase the physiological situation is more complex owing to the presence of bile salts in the intestine. These amphipathic salts, and phospholipids or proteins, prevent the binding of pancreatic lipase to the lipid–water interface¹¹. Therefore pancreatic lipase needs a small protein partner, colipase, also secreted by the exocrine pancreas, which allows the enzyme to bind to the bile salt-covered interface^{12,13}. The three-dimensional structure of the lipase–procolipase complex illustrates how this might happen: procolipase binds to the non-catalytic β -sheet of the C-terminal domain of the lipase and exposes the hydrophobic tips of its fingers at the opposite site of its lipase-binding domain¹⁴. This hydrophobic surface helps to bring the catalytic domain of lipase into close contact with the interface.

Here we determine the three-dimensional structure of the procolipase–lipase complex in the presence of mixed phosphatidylcholine/bile salt micelles at 3.0 Å resolution. In this

structure, pancreatic lipase reorganizes its lid and has a phospholipid bound in the active site. Important conformational changes take place at residues lining the lid, finally creating a hydrophobic active-site groove and adjusting the hydrolytic machinery.

The importance of the lid in mammalian lipases has been illustrated by mutagenesis of lipoprotein lipase^{15,16}. This enzyme shows considerable sequence homology to pancreatic lipase and hepatic lipase. These three lipases belong to the same family and probably have the same overall three-dimensional structure¹⁷. Sequence comparison reveals the presence of a lid bounded by two conserved cysteines in pancreatic, hepatic and lipoprotein lipase. The three lids are of comparable length, but sequence homology between pancreatic lipase on the one hand and hepatic and lipoprotein lipase on the other is poor. The crystal structure presented here may explain why a chimaeric construction consisting of the pancreatic lipase lid and the lipoprotein lipase core yields a completely inactive enzyme: the reorganized pancreatic lid may not encounter the correct stabilizing interactions with the core of the protein when grafted onto the lipoprotein lipase.

Resolution of the structure

The crystallization and structure determination of the human pancreatic lipase–procolipase complex have been described^{14,18}. Co-crystallization of this complex with mixed micelles of 1,2-didodecanoyl-*sn*-3-glycerophosphorylcholine and sodium taurodeoxycholate resulted in crystals of a different spacegroup ($P4_2, 2_1$; $a = b = 133.4$ Å, $c = 92.6$ Å)¹⁸. First, a concentrated solution of mixed micelles was prepared in water (30 mM phosphatidylcholine and 60 mM sodium taurodeoxycholate). One microlitre of this micellar solution was added to a pre-equilibrated sitting drop of 4 μ l of a solution of the complex in water (8 mg ml^{−1} pancreatic lipase and 2.5 mg ml^{−1} procolipase) and 4 μ l of the well solution containing 0.1 M MES, pH6.0, 0.3 M NaCl and 2% polyethylene glycol PEG8000. Two days later, crystallization was started by adding 1 μ l 250 mM β -octylglucoside to this drop. These crystals are very difficult to obtain and remain small (0.05 mm $\times$ 0.05 mm $\times$ 0.100 mm). Despite their size, they diffract to 3.0 Å resolution. The structure has been solved with data collected from a single crystal using the molecular replacement option of X-PLOR¹⁹ (for details, see Fig. 1 legend). The electron density of a fragment of the molecule

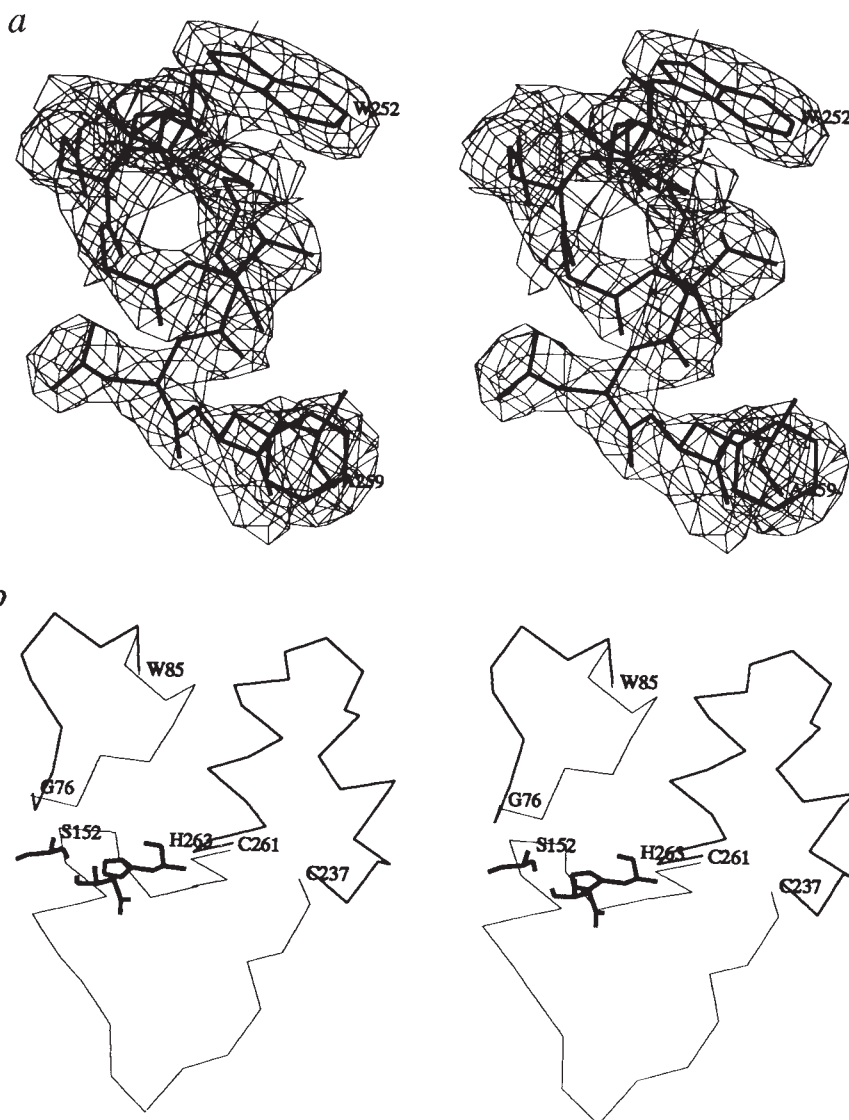
FIG. 1 *a*, Final electron density map using coefficients ($2F_o - F_c$) and calculated phases (F_o and F_c are the observed and calculated structure factors) from the refined model of the open structure of the pancreatic lipase-procolipase complex. The main helix (residues 252 to 259) of the rearranged lid is represented. *b*, Detailed stereographic view of the conformational change of the $\beta 5$ loop (residues 76 to 85) and the lid (residues 237 to 261) ($C\alpha$ atoms). The closed and open structures are represented by thin and thick lines, respectively. The catalytic triad residues are also shown: Ser 152, His 263 and Asp 176 (no label). The disulphide bridge (Cys 237-Cys 261) remains fixed and anchors the lid to the protein core.

METHODS. Native data were collected on a small single crystal, using a Rigaku RU-200 rotating anode X-ray source and a Mar Research Imaging plate detector (Hamburg, Germany). We measured 97,990 reflections that eventually yielded 15,496 independent reflections between 8.0 Å and 3.0 Å resolution (90% completion; $R_{\text{sym}} = 20\%$ up to 3.0 Å). Data was reduced using MOSFLM (A. Leslie) and the CCP4 suite of programmes (Daresbury, England). The structure was solved by molecular replacement using XPLOR¹⁹. As a search model we used only the refined coordinates of human pancreatic lipase. Inspection of the packing of the molecular replacement solution revealed a clash between the lid regions (residues 239 to 260) of symmetry-related molecules. This oriented model (without the lid) was submitted to simulated annealing refinement (X-PLOR). A $2F_o - F_c$ map revealed that a second region in the lipase molecule, comprising residues 76 to 85, underwent an important conformational change. The complete lipase model was further refined and a new $2F_o - F_c$ map revealed the presence of the procolipase molecule. As parts of the procolipase structure also changed conformation, we preferred partially to reconstruct the procolipase molecule into the electron density. After refinement of the complex (R factor = 19.3%), there was evidence in the residual electron density for bound β -octylglucoside (mandatory for crystallization) at the entrance of the active site and for the presence of a phospholipid in the active site. β -Octylglucoside could readily be placed into the electron density but the precise conformation of the lipid in the active site was more equivocal. Density was clear for two bound acyl chains, but the position of the phosphorylcholine headgroup and the orientation of the lipid are tentative. The final crystallographic R factor of 0.186 is for all reflections between 8.0 and 3.0 Å resolution. The final model, containing 4,132 protein atoms, one Ca^{2+} ion,

is shown in Fig. 1*a*. Coordinates have been submitted to the Brookhaven Protein Data Bank.

Conformational changes of the complex

The induction of a conformational change upon binding of pancreatic lipase to the lipid-water interface was anticipated long before the crystal structure of the enzyme was known²⁰. It was obvious that the pancreatic lipase lid had to move away in order to give the substrate access to the active site. A blocking α -helix (residues 248 to 255) is connected to the body of the molecule by two elongated polypeptide stretches (residues 238 to 247 and residues 256 to 260) that join through a disulphide bridge (Cys 237-Cys 261) before continuing into the body of the catalytic domain (Fig. 2*a*). The presence of mixed micelles in the crystallization solution has a marked influence on the structure of the pancreatic lipase-procolipase complex. The structures determined in the presence and absence of mixed micelles will hereafter be called open and closed, respectively. The opening of the active site is caused by a structurally complicated reorganization of the complete region contained between the two cysteines mentioned. The first three residues after

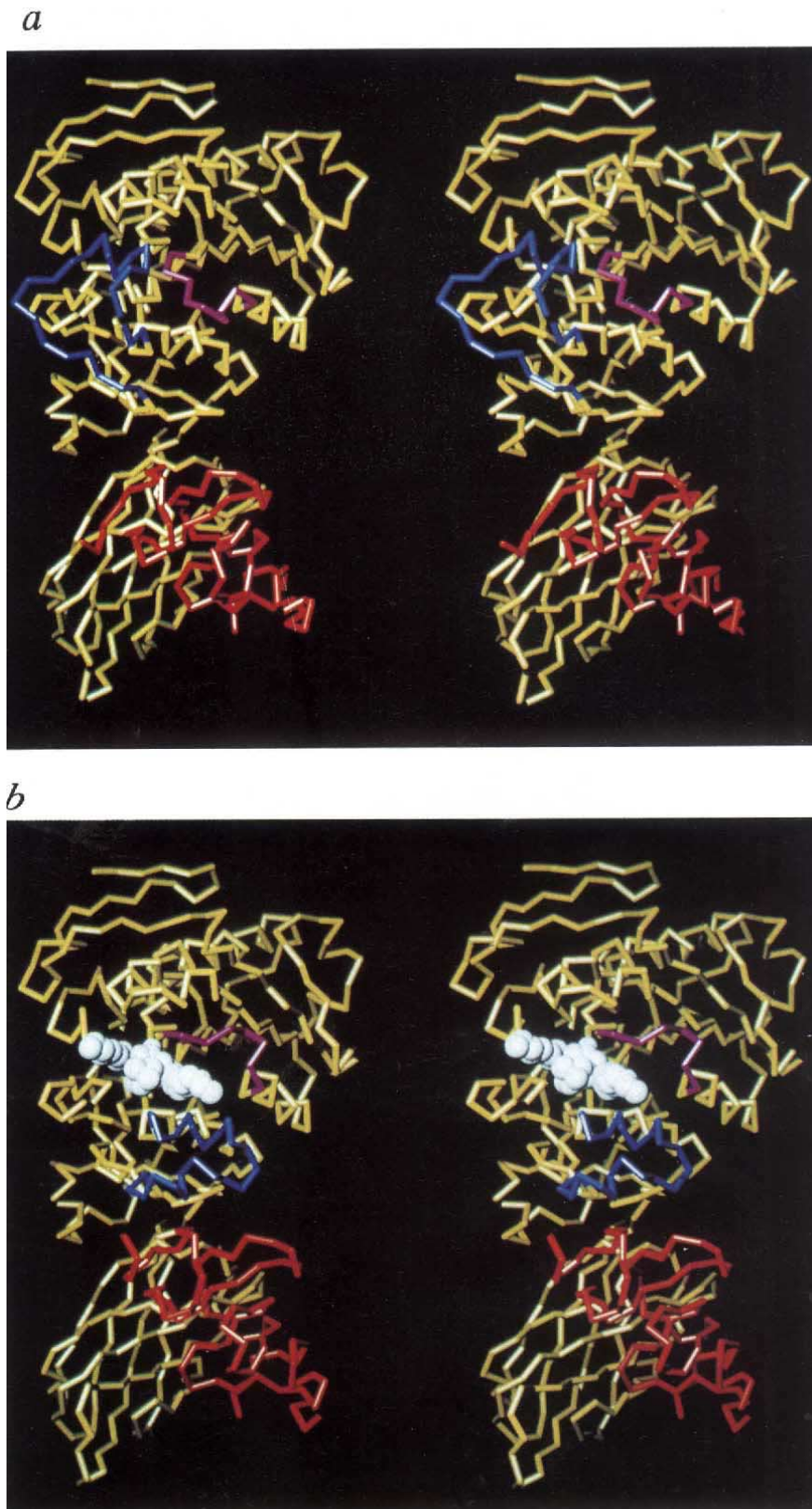


one phosphatidylcholine and one β -octylglucoside molecule, has good stereochemistry: the r.m.s. deviations from ideal values for bonding distances and bond angles are 0.016 Å and 3.6°, respectively. All model construction and structure comparison work used the program TURBO, implemented on a Silicon Graphics 4D/380 computer²⁷. $R_{\text{sym}} = \sum |I(I)| / \sum \langle I \rangle$.

Cys 237 remain in extended conformation, but the direction of the strand changes drastically as a result of the different main-chain dihedral angles of Lys 238. The flap contains α -helices in the open and closed structures but they are formed by different residues: the helix of the closed conformation (residues 248 to 255) partially unwinds and two new helices are formed (residues 241 to 246, and residues 251 to 259). Residues 241 to 246, which are extended in the closed lid, are wound up into a short irregular helix and are followed by a tight turn comprising residues 247 to 250 in the open form (Fig. 1*b*). The remaining part of the open lid, from residue 251 to 259, forms a regular helix. Residues 256 to 260 are in extended conformation in the closed form. Ala 260, which has a β -conformation in the open form, connects the main helix of the open lid to a helix of the α/β -fold of the catalytic domain. Both helices are at an angle of $\sim 90^\circ$. The change in conformation of the lid results in a maximal main-chain movement of 29 Å for Ile 248. The open lid has the shape of an elbow that sticks out from the core of the catalytic N-terminal domain (Fig. 2*b*).

The functional consequences of the structural reorganization of the lid are as follows. (1) The substrate has access to the

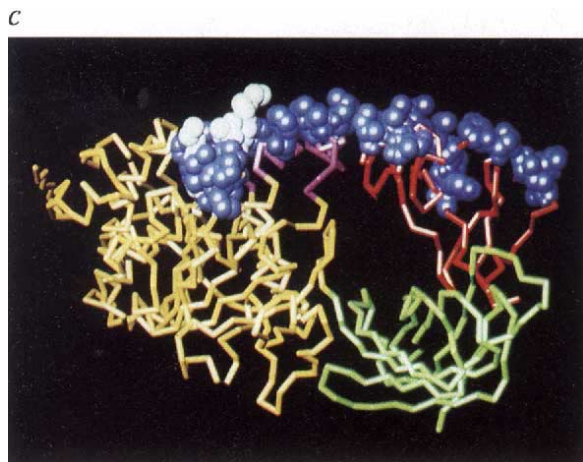
FIG. 2 *a* and *b*, Stereo pictures of the comparison of the closed (*a*) and open (*b*) forms of the lipase-procolipase complex ($C\alpha$ atoms). Pancreatic lipase is represented in yellow. The lid (residues 237 to 261) is shown in blue and the $\beta 5$ loop (residues 76 to 85) in mauve; colipase is in red. The open form has also a bound phospholipid in the active site (white spheres). The picture shows how the active site, hidden under these two surface loops opens up like the petals of a flower. The conformational transitions in the lid and in the $\beta 5$ loop are characterized by a profound change in secondary and tertiary structure. *c*, Hypothetical lipid-binding surface of the lipase-procolipase complex. The lipase catalytic N-terminal domain is in yellow, the C-terminal domain in green and colipase is in red. Hydrophobic residues are represented as blue spheres and the model of the bound phospholipid is in white. The change in conformation of the lid brings its own hydrophobic residues to the surface and uncovers many others around the active site. The open lid together with the hydrophobic tips and N-terminal region of procolipase form a continuous hydrophobic plateau that extends over more than 50 Å.



active site: Trp 252 particularly, which fills the active-site pocket in the closed structure, is conveyed to the surface of the molecule; (2) a second loop, which leans against the lid when it is closed and makes few contacts with the rest of the protein, now changes its conformation and creates the oxyanion hole (see later); (3) an important hydrophobic surface that spreads over the entrance of the active site is created; (4) the catalytic N-terminal domain of lipase interacts with procolipase via the open lid.

The movement of Trp 252 is supported by fluorimetric measurements, in which a large change in fluorescence is observed upon acylation of the active serine by tetrahydrolipstatin²¹. From these experiments it was concluded that lipase undergoes a conformational transition upon inhibitor binding, resulting in a change of the micro-environment of the tryptophans.

The restructuring of the lid causes an important conformational change in the active-site loop which is situated between strand $\beta 5$ and helix $\alpha 2$ of the catalytic domain and contains



residues 76 to 85 (secondary structure nomenclature after ref. 1). This loop (β -loop) is making van der Waals contacts exclusively with the closed lid. Without this support, owing to the movement of the lid, the loop folds back upon the core of the protein (Figs 1b, 2b). This results in main-chain movements from 2 Å for Phe 77 to 13 Å for Glu 83 (Figs 1b, 2a and b). The β 5 loop changes drastically in direction at Ile 78, whose ϕ , ψ angles shift from the β -sheet region (-118° , 124°) to an energetically unfavourable region (68° , -37°), and at Glu 83, which is α -helical in the closed structure (-50° , -40°) and which now adopts a β -sheet conformation (-67° , 148°). The new conformation of the β 5 loop can be described as a succession of two β -turns (residues 76 to 79, and 80 to 83). The side chain of Phe 77 occupies the space created by the removal of Trp 252.

The closed lid is making contacts that are almost exclusively non-polar with the rest of the protein. The open lid, however, is stabilized by many new hydrogen bonds and salt bridges, partly with the core of the protein, partly with the β 5 loop, and partly with the N-terminal region of procolipase (Fig. 3). For instance, Arg 256 makes a salt bridge with Asp 79 (β 5 loop) and a hydrogen bond with the Tyr 267 hydroxyl in the core of the N-terminal domain. As in the case of the lid, the new conformation of the β 5 loop is stabilized by many new polar interactions with amino acids from the core of the molecule and amino acids from the lid. The concerted conformational changes of the lid and of the β 5 loop increase considerably the hydrophobicity of the region around the active site. This is achieved by exposing hydrophobic side chains as well as by burying hydrophilic ones. An important contribution to this hydrophobic surface comes from the hydrophobic side of the amphipathic main helix of the lid.

The lid in pancreatic lipase and in *Rhizomucor miehei* lipase originate from topologically different secondary structure ele-

ments. They are therefore approaching the catalytic triad from opposite directions, although a tryptophan is blocking the active site in both structures. The structural rearrangements of the lid in pancreatic lipase are very different from those in *Rhizomucor miehei* lipase^{9,10}: upon inactivation of this latter enzyme with a covalent inhibitor directed against the active site serine, the amphipathic lid-helix rolls back upon the core of the protein and brings its hydrophobic surface to the outside. This conformational change consists of a rigid body movement of the helical lid around two hinge regions. Most of the main and side chains keep their original conformation, in contrast to the lid opening in the pancreatic lipase-procolipase complex.

Formation of the active site

The position and conformation of amino acids forming the catalytic triad are not significantly different in the open and closed structures. The combined conformational changes of the lid and the β 5 loop, however, have changed the environment of the catalytic triad dramatically. The active site serine has become totally accessible to solvent and is lying at the bottom of a hydrophobic canyon which is well adapted for the binding of a lipid substrate.

It is known that the functioning of the catalytic triad of serine proteases is assisted by the presence of an oxyanion hole, an electrophilic region that stabilizes the negative charge generated during nucleophilic attack of the carbonyl carbon of the substrate ester bond by the Ser O γ ²². The importance of this transition state stabilization has been well established in the case of subtilisin by site-directed mutagenesis²³. The architecture of the oxyanion hole and the contribution of oxyanion hole stabilization to catalysis, although predicted to be part of the catalytic mechanisms, is less well documented in the case of lipases and esterases. It has been suggested from modelling studies on the interaction of pancreatic lipase with a long-chain acyl substrate that the oxyanion hole could be formed by the main-chain nitrogens of Leu 153 and Phe 77²⁴. To fulfil this function, Phe 77 must move aside to provide space for the binding of the tetrahedral intermediate. This prediction is confirmed by our crystal structure: the conformational transition of the β 5 loop removes the main-chain nitrogen of Phe 77 2 Å farther away from the Ser 152 O γ oxygen and brings the nitrogen into an ideal position to stabilize the negative charge developed during ester hydrolysis, at 5.2 Å from the O γ of Ser 152 (Fig. 4b). The side chain χ 1 dihedral angle of Phe 77 shifts from 43° to -170° .

Modelling of lipid into active-site density

After refinement of the protein, there still remained an important $F_o - F_c$ residual electron density in the active site. This density consists of a long and continuous tube, starting at the entrance of the active site cleft, touching the Ser 152 O γ oxygen and continuing on its way out. As the crystals were grown in the

FIG. 3 Stereographic picture of the interactions between the lipase lid, the rest of the protein and colipase. The open conformation of the lid (indicated by residues W252, D257, R256 and N240) is stabilized by new interactions with the β 5 loop (E83), the core of the protein (Y267, K268) as well as with procolipase (E15#, R38#). Some of these hydrogen bonds are represented by dashed lines.

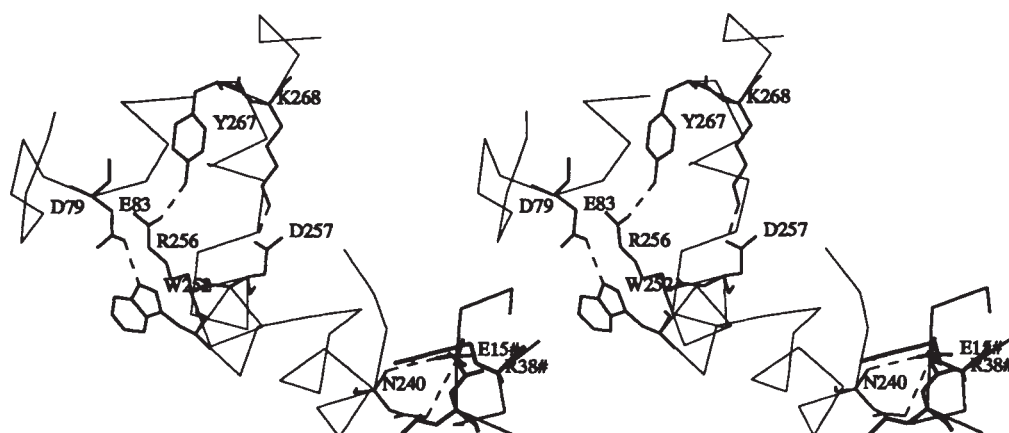
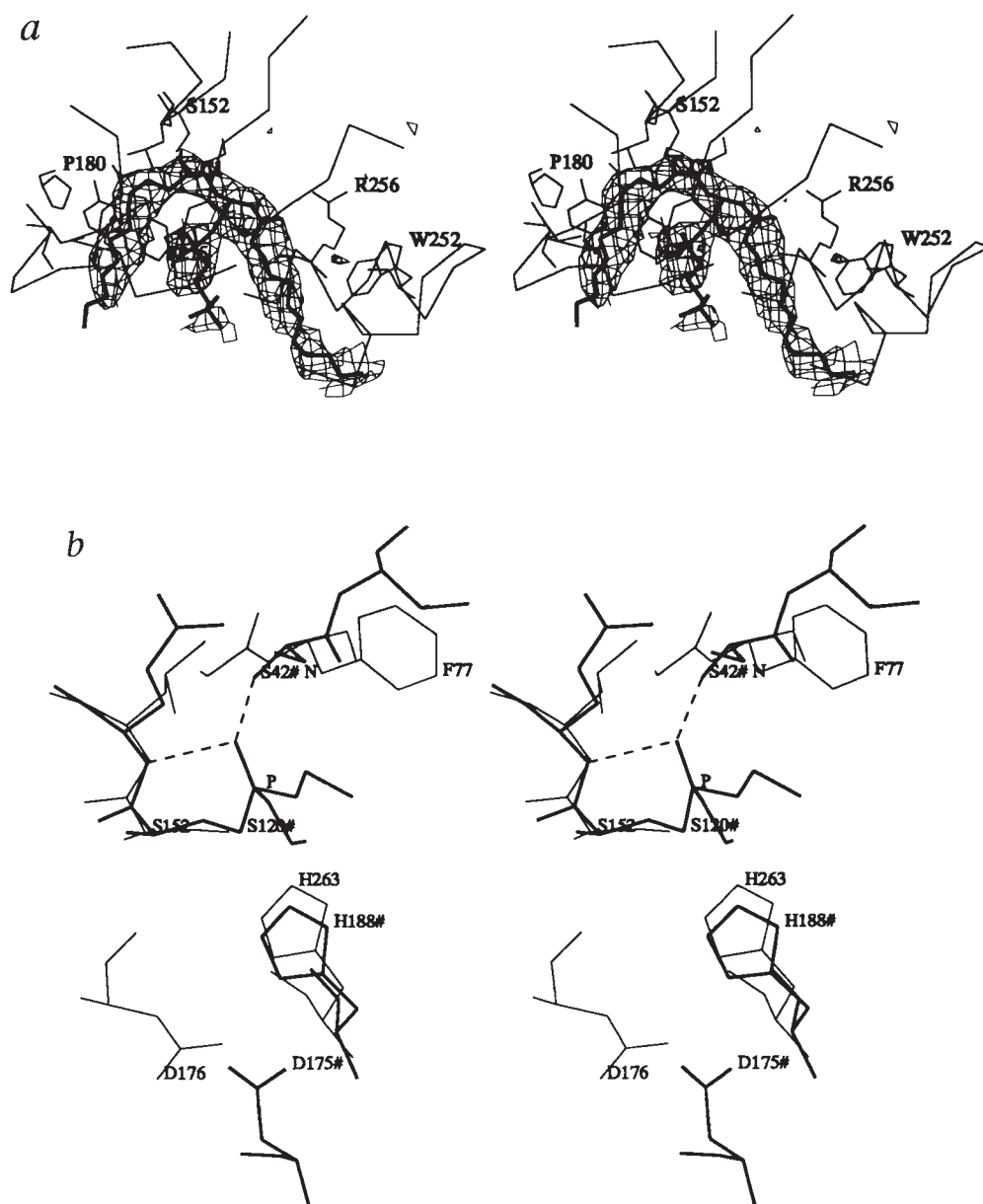


FIG. 4 *a*, Stereographic view of the electron density difference map at the 2σ level in the active site, using coefficients $F_o - F_c$ and calculated phases. F_c values are calculated structure factors after refinement of the lipase–procolipase complex. Thick lines represent the refined model of a bound 1,2-didodecanoyl-*sn*-3-glycerophosphorylcholine molecule. The active-site serine and part of the hydrophobic residues of the active-site groove are represented by thin lines. C1 marks the equivalent of the carbonyl bond that is hydrolysed in triglycerides. The C2 carbonyl oxygen in this model is making a hydrogen bond to the NE nitrogen of Arg 256. *b*, Comparison of the oxyanion holes of pancreatic lipase in its open form (thin lines) and cutinase (thick lines), which has the inactivator diethylphosphate (P) covalently bound to its catalytic Ser 120. Cutinase residues are indicated by hash symbols. The native structure was determined at 1.6 Å resolution²⁸ and the structure of the enzyme covalently inactivated by reacting the catalytic serine with diethyl-*p*-nitrophenylphosphate (E600) at 1.9 Å (C. M. unpublished results). The catalytic triad atoms and some of the well conserved central β -sheet C α atoms were used for superposition of the two structures (rigid body least-squares fitting option in TURBO). The main-chain nitrogens of residues Asn 121 and Ser 42 in cutinase and residues Leu 153 and Phe 77 in lipase are perfectly positioned to stabilize the oxyanion generated during nucleophilic attack of the carbonyl group of the substrate by the Ser O γ . The nitrogen of Phe 77 moves to this position by the conformational transition of the β 5 loop. The C1 carbonyl group of the phospholipid as shown in *a* is superimposed on the phosphoryl group of E600.



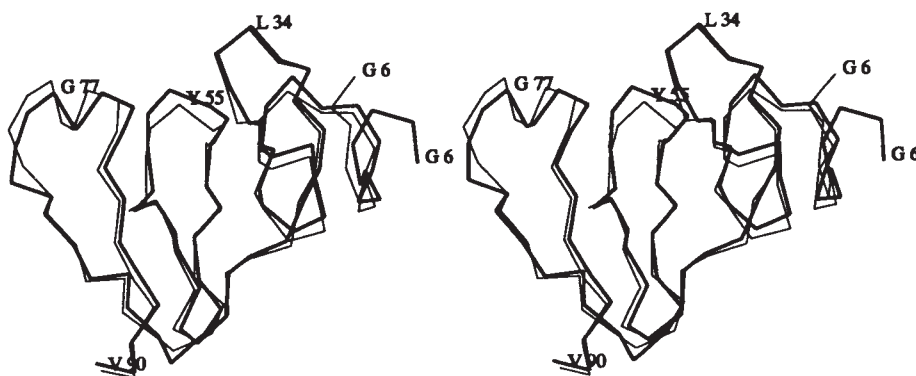
presence of mixed micelles of taurodeoxycholate and phosphatidylcholine, we modelled a phosphatidylcholine molecule into the residual electron density, but the medium resolution of the map precludes unequivocal interpretation.

The residual electron density could best be accommodated by a phospholipid in the configuration shown in Fig. 4*a*. The $2F_o - F_c$ electron density for the acyl chains is at the 2σ level, and the polar parts of the ester atoms are at 1σ . No density was observed for the quaternary ammonium group of the choline moiety. The two dodecanoyl chains could easily be placed into the electron density left and right of Ser 152. The *sn*-1 acyl chain is passing through a hydrophobic channel formed by aromatic side chains (Tyr 114, Phe 215, Phe 77) and by Pro 180, Ile 209 and Leu 202. The side chain of Phe 215 is rotated around its χ_1 angle over 115° relative to the closed structure and is positioned to hold the *sn*-1 acyl chain in a sandwich with Tyr 114. The *sn*-2 acyl chain is lying in a hydrophobic groove, formed by residue side chains from the main helix of the flap region (residues 251 to 259) and by Ile 78, coming from the β 5 loop. The end of the acyl chain is wrapped around Trp 252. Part of the chain is lying

along the side chain of Arg 256. Together the two acyl chains completely cover the hydrophobic active-site canyon.

The carbonyl carbon of the primary ester bonds, which is scissile in triglyceride substrates, can be positioned in front of the Ser 152 O γ oxygen (Fig. 4*a*). The carbonyl oxygen is pointing towards the main-chain nitrogens of Phe 77 (at 3.1 Å) and Leu 153 (at 2.7 Å), further confirming their role in the formation of the oxyanion hole. Of course in this configuration the phospholipid should be hydrolysed. We therefore cannot exclude that it is actually a mixture of lysophosphatidylcholine and fatty acid that is bound to the active site. Although the relative level of phospholipase activity of pancreatic lipase is controversial, there is a general agreement that this activity of pancreatic lipase is marginal, compared with its activity towards triglycerides¹¹. The crystallization conditions (high enzyme concentration and long incubation times) could have led to some hydrolysis of the phospholipids in the drop. Cutinase, an esterase with a catalytic triad that is not hidden by a surface lid, also hydrolyses triglycerides but does not show interfacial activation. We have superimposed the catalytic triads of covalently inhibited cutinase and

FIG. 5 C α coordinate superposition of the procolipase structures as present in the open (thick lines) and closed (thin lines) lipase-procolipase complex. The largest differences are seen in the N-terminal region. The peptide containing residues 6 to 10 is stabilized in the closed complex by crystallographic contacts and is in the open form involved in interactions with the lid.



of pancreatic lipase with bound phospholipid. In Fig. 4b the oxyanion hole is shown to be preformed in cutinase, whereas in pancreatic lipase it is created by interaction of the enzyme with the substrate. A substrate-induced fit may also arise in *Rhizomucor* lipase, in which the side chain of Ser 82 rotates upon binding of an active-site-directed inhibitor and participates in the formation of the oxyanion hole¹⁰.

At the entrance of the active site of the open form, facing Trp 252, there remained an important $2F_o - F_c$ electron density after refinement (at the 3σ level), which could clearly be attributed to a bound β -octylglucoside molecule (not shown). β -Octylglucoside is a necessary component of the crystallization mixture¹⁴. The packing of the molecules in the crystal brings the active sites of two symmetry-related molecules face to face, each with its bound phospholipid and detergent molecule. We suggest that this extended hydrophobic environment is mimicking a lipid-water interface and that therefore the observed crystal structure is relevant to what is happening to the pancreatic lipase on penetration into the lipid phase.

Procolipase interacts with the open lid

The interactions between procolipase and the lipase C-terminal domain are preserved in the presence of lipids. Binding is still conferred by two hairpin loops and the C-terminal region of the procolipase and by residues from the central β -sheet of the C-terminal domain of lipase¹⁴. Besides these interactions, procolipase is also binding to the lid region. This interaction is mediated by three direct hydrogen bonds (Fig. 3). The interacting residues of the lid (Val 246, Ser 243 and Asn 240) and of procolipase (Arg 38, Leu 16 and Glu 15) are strictly conserved in all known pancreatic lipase and procolipase sequences. The side-chain carboxylate of Glu 15 of procolipase forms a hydrogen bond with the main-chain nitrogen of Asn 240 of the lipase. Chemical modification of this Glu 15 resulted in a complete loss of the procolipase activity, suggesting that the interaction between the lid and procolipase is of prime importance for the functioning of the complex²⁵ (L. Sarda, unpublished results).

Three structural rearrangements are responsible for the new contact between the N-terminal domain and colipase: (1) the lid movement, as described, brings it towards procolipase; (2) conformational changes in the N-terminal part of procolipase; (3) a rigid bending motion of the C-terminal domain of lipase (residues 337 to 449) relative to the N-terminal domain also results in the approach of procolipase to this domain.

Procolipase consists of three fingers, held together by disulphide bridges, and a small N-terminal region¹⁴. The structure of the procolipase is well defined in the proximity of the five disulphide bridges that are running across the molecule, but the tips of the fingers are very mobile. As we found previously, the five N-terminal and the two C-terminal amino acids of procolipase are not detectable. Superposition of the procolipase

structures in the open and closed lipase-procolipase complexes shows that they are very similar (Fig. 5), except in the N-terminal region (residues 6 to 25), where there is a systematic main-chain atom shift of ~ 2 Å and for some residues in the poorly defined tips of the fingers. The differences in procolipase structure are considerable in the N-terminal peptide (residues 6 to 9), which in the lipid-free complex adopts an elongated conformation stabilized by a crystallographic symmetry contact. In the present open structure this peptide forms a hook, exposing three successive isoleucines (Ile 7 to 9) in the same direction as the fingers. Peptide cleavage shows that these isoleucines are important for the function of procolipase²⁶. This N-terminal conformation may be stabilized by interaction with the open flap.

The third reason for the interaction between procolipase and the flap is a bending motion of the C-terminal domain of lipase towards the catalytic N-terminal domain. This can best be described as a rigid rotation of about 7° around a hinge (residues 334 to 336) situated at the contact region between the catalytic N-terminal domain and the C-terminal domain. This bending results in movements of 5 Å for residues at the extremity of the C-terminal domain but, most importantly, it brings procolipase much closer to the catalytic domain of lipase.

Some variation has been observed in the bending angle between the N-terminal and C-terminal domains when the structures of two crystallographic copies of the human pancreatic lipase were compared. We have evidence from the crystal structure of horse pancreatic lipase that this bending angle can be influenced by crystallographic contacts and is not solely determined by the presence or absence of procolipase and/or lipids (Y. Bourne, unpublished results).

We have emphasized the overall amphipathic character of the procolipase molecule: most of the hydrophobic residues are located at the tips of the fingers, whereas hydrophilic residues are found at the opposite site and form the binding site for the C-terminal domain of lipase¹⁴. The hypothesis that the tips of the fingers constitute the lipid-binding site is confirmed by the present structure determined in the presence of mixed micelles. The open lid and the extremities of the procolipase fingers form an impressive continuous hydrophobic plateau, extending over more than 50 Å (Fig. 2c). This surface must be capable of interacting strongly with lipid/water interfaces covered with bile salts, a task that pancreatic lipase without procolipase is not able to accomplish.

The present crystal structure shows that the pancreatic lipase lid is a complicated structural device. Upon lipid binding it reorganizes in an ingenious way. This conformational change serves many purposes: it creates the interfacial lipid-binding site, gives shape and access to the active-site canyon, stabilizes the oxyanion hole loop, and strengthens the interaction with procolipase. It may well be that the lid regulates the different substrate specificity, interfacial behaviour and cofactor dependency of mammalian lipases. □

Received 21 December 1992; accepted 22 March 1993.

1. Winkler, F. K., D'Arcy, A. & Hunziker, W. *Nature* **343**, 771–774 (1990).
2. Brzozowski, A. M. *et al. Nature* **351**, 491–494 (1991).
3. Schrag, J. D., Li, Y., Wu, S. & Cygler, M. *Nature* **351**, 761–764 (1991).
4. Ollis, D. V. *et al. Prot. Engng* **5**, 197–211 (1992).
5. Lowe, M. E. *J. biol. Chem.* **267**, 17069–17073 (1992).
6. Emmerich, J. *et al. J. biol. Chem.* **267**, 4161–4165 (1992).
7. Faustinella, F., Smith, L. C., Semerkovich, C. F. & Chan, L. *J. biol. Chem.* **266**, 9481–9485 (1991).
8. Faustinella, F. *et al. J. biol. Chem.* **266**, 14418–14424 (1991).
9. Brzozowski, A. M. *et al. Nature* **351**, 491–494 (1991).
10. Derewenda, U., Brzozowski, A. M., Lawson, D. M. & Derewenda, Z. S. *Biochemistry* **31**, 1532–1541 (1992).
11. Verger, R. in *Lipases* (eds Borgström, B. & Brockman, H.) 83–150 (Elsevier, Amsterdam, 1984).
12. Sarda, L. in *Lipases: Structure, Mechanism and Genetic Engineering* 145–155 (GBF Monographs **16**, 1990).
13. Borgström, B. & Erlanson, C. in *Lipases* 151–183 (Elsevier, 1984).
14. van Tilbeurgh, H., Sarda, L., Verger, R. & Cambillau, C. *Nature* **359**, 159–162 (1992).
15. Faustinella, F., Smith, L. C. & Chan, L. *Biochemistry* **31**, 7219–7223 (1992).
16. Dugi, K. A., Dugi, K. A., Dichek, H. L., Talley, G. D., Brewer, H. B. Jr & Santamarina-Fojo, S. *J. biol. Chem.* **267**, 25086–25091 (1992).

17. Kirchgeessner, T. G. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 9647–9651 (1989).
18. van Tilbeurgh, H. *et al. J. molec. Biol.* **229**, 532–554 (1993).
19. Brünger, A. T. *Acta crystallogr.* **46**, 46–57 (1990).
20. Desnuelle, P., Sarda, L. & Ailhaud, G. *Biochem. biophys. Acta* **37**, 570–571 (1960).
21. Lüthi-Peng, Q. & Winkler, F. K. *Eur. J. Biochem.* **205**, 383–390 (1992).
22. Krau, J. A. *Rev. Biochem.* **46**, 331–358 (1977).
23. Bryan, P., Pantoliano, W., Quill, S. G., Hsiao, H.-Y. & Poulos, T. *Proc. natn. Acad. Sci. U.S.A.* **83**, 3743–3745 (1986).
24. Gubernator, K., Müller, K. & Winkler, F. K. in *Lipases: Structure, Mechanism and Genetic Engineering* (eds Alberghina, L., Schmid, R. D. & Verger, R.) 9–16 (Weinheim VCH, New York, 1991).
25. Erlanson-Albertsson, C. *Biochim. biophys. Acta* **1125**, 1–7 (1992).
26. Erlanson-Albertsson, C. & Larsson, A. *Biochim. biophys. Acta* **665**, 250–255 (1981).
27. Roussel, A. & Cambillau, C. in *Silicon Graphics Directory* (Silicon Graphics, Mountain View, California, 1991).
28. Martinez, C., de Geus, P., Lauwereys, M., Mathysens, G. & Cambillau, C. *Nature* **356**, 615–618 (1992).

ACKNOWLEDGEMENTS. We thank Y. Gargouri for lipase purification, L. Sarda for discussion, and F. Carriere for mixed lipid micelles. This project was supported by the EEC BRIDGE T-lipase-programme, the CNRS IMABIO programme and the PACA region.

LETTERS TO NATURE

Hard X-rays from accretion disk boundary layers

Ramesh Narayan & Robert Popham

Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138, USA

ACCRETION disks^{1,2} are found in many astrophysical objects, ranging from newly formed stars and mass-transferring binary systems to quasars and other active galactic nuclei. An important feature of accretion disks is the boundary layer—the interface between the disk and the accreting objects—where up to half the accretion luminosity may be liberated. The lack of a satisfactory description of the flow and thermal structure of this layer has long been a handicap when modelling disk spectra. Here we report numerical solutions of a model of thin accretion disks around a central white dwarf which includes a self-consistent description of the boundary layer. We find two distinct kinds of solution depending on the mass accretion rate $\dot{M}$. At high rates, we find optically thick boundary layers whose radial width and peak temperature decrease with decreasing $\dot{M}$, but when the accretion rate falls below a critical value, the boundary layer becomes optically thin, and the width and temperature increase dramatically. Our results provide an explanation for the hard X-rays observed³ in cataclysmic variables, particularly at low $\dot{M}$. It should be possible to extend our analysis to other accretion-disk systems.

We consider an accretion disk around a central white dwarf of mass $M_* = 1$ solar mass ($M_\odot$) and radius $R_* = 5 \times 10^8$ cm. Solutions were obtained by numerically solving a set of coupled differential equations which describe the hydrodynamical and thermal structure of a steady-state thin accretion disk as a function of radius R . The radial and tangential components of the momentum equation give^{4,5}

$$v \frac{dv}{dR} - (\Omega^2 - \Omega_K^2) R + \frac{1}{\rho} \frac{dP}{dR} = 0 \quad (1)$$

$$\Omega_K(R) = \left(\frac{GM_*}{R^3} \right)^{1/2}$$

$$v \frac{d\Omega}{dR} - v \left(\Omega - \frac{J}{R^2} \right) = 0 \quad (2)$$

where v is the radial velocity at radius R , Ω is the angular velocity, Ω_K is the keplerian value of Ω , ρ is the density, P is the pressure, ν is the coefficient of shear viscosity, and J is an eigenvalue describing the inward flux of angular momentum per unit mass. We use a modified prescription^{6,7} for ν , which is an extension of the usual α -prescription of Shakura and Sunyaev¹, designed to satisfy causality. We set $\alpha = 0.1$.

A crucial feature of this work is the inclusion of an energy equation^{8–10}

$$\dot{M} T \frac{dS}{dR} - \dot{M} \frac{\nu}{v} \left(\frac{d\Omega}{dR} \right)^2 - \frac{d}{dR} (4\pi R H F_R) - 4\pi R F_V = 0 \quad (3)$$

where the four terms correspond to entropy advection, viscous energy generation, divergence of radial radiative flux F_R , and energy loss due to vertical flux F_V through the top and bottom of the disk; H is the half-thickness of the disk, such that $\dot{M} = 4\pi R H \rho |v|$. We assume an ideal equation of state, corresponding to a perfect gas mixed with radiation.

For the vertical flux F_V we use approximate solutions¹¹ relating the effective temperature T_{eff} , defined by $F_V \equiv \sigma T_{\text{eff}}^4$, to the central temperature T . We use a simple grey opacity which includes Kramers free-free absorption and electron scattering. For the radial flux F_R , we use a 'two-stream' model of radiative transfer, which leads to two differential equations for the radiation density U and radial flux F_R . We thus have a total of five first-order ordinary differential equations. We solve these numerically with three inner boundary conditions at $R = R_*$: these are $\Omega = \Omega_*$, $P = P_*$, $F_R = \sigma (T_*)_{\text{eff}}^4$, where we arbitrarily assume $(T_*)_{\text{eff}} = 2 \times 10^4$ K; and two outer boundary conditions at $R = R_{\text{out}} = 100 R_*$: these are $\Omega = \Omega_K$, $U = a T^4$. We use a relaxation method to compute the solutions, taking care to maintain sufficient resolution in the difficult boundary layer region.

Figure 1 shows typical solutions at high $\dot{M}$ for a star rotating with angular frequency $\Omega_* = 0.5 \Omega_b$, where $\Omega_b \approx \Omega_K(R_*)$ is the 'break-up limit' on the spin of the star. The boundary layers in these solutions are optically thick in the vertical direction, and they have two distinct radial widths. The dynamical width over which Ω changes from Ω_K to Ω_* is narrow, less than $0.01 R_*$, whereas the thermal width over which the boundary-layer luminosity is radiated is larger, $\sim 0.1 R_*$. Both widths decrease monotonically with decreasing $\dot{M}$, as do the central temperature T and effective temperature T_{eff} . Further, the boundary-layer luminosity decreases as Ω_* is increased, vanishing when the star reaches break-up ($\Omega_* = \Omega_b$). None of these results is surprising, but self-consistent solutions have not been previously calculated for this region of a disk.

Figure 2 shows some of our solutions at low accretion rates, where the boundary layer is optically thin. We see immediately very different behaviour from that in Fig. 1. Now, as $\dot{M}$ decreases, both the width and the temperature of the boundary layer increase very rapidly. Indeed, for $\dot{M} = 10^{-10.5} M_\odot \text{ yr}^{-1}$ and $\Omega_* = 0.1 \Omega_b$, the width ΔR_{BL} of the boundary layer is comparable to the stellar radius R_* and the maximum temperature T_{BL} is well in excess of 10^8 K. Both ΔR_{BL} and T_{BL} increase even further for lower $\dot{M}$, but decrease as Ω_* increases. Further, the results are sensitive to the choice of α . For $\dot{M} = 10^{-9.5} M_\odot \text{ yr}^{-1}$ and

$\Omega_* = 0.1\Omega_b$, as we increase α from 0.1 to 0.4, we find that ΔR_{BL} increases from $0.13R_*$ to $0.43R_*$ and T_{BL} goes from 10^8 K to 2×10^8 K. At these extreme temperatures the disk is no longer very thin; nevertheless, we expect the models to be reasonably realistic.

The surprising behaviour of the boundary layer at low $\dot{M}$ can be explained in terms of a thermal instability^{3,12}. The accreting gas cannot cool efficiently because of the decreased density in the optically thin boundary layer. This forces the disk to store most of the viscously generated energy in the form of entropy, heating the boundary-layer gas rapidly. The gas then expands vertically, which serves to decrease its density further. Thus in, much of the boundary-layer region, we do not have a balance between viscous heating and locally radiated vertical flux (the second and fourth terms in equation (3)), as in the rest of the disk. Instead, viscous heating is balanced primarily by entropy advection, the first term in equation (3), and the two radiative flux terms play a relatively small role. When the superheated material in the boundary layer finally reaches the star, the thermal instability works in reverse. The fluid now becomes optically thick and cools abruptly. Although there is no hydrodynamic shock at this point⁷, the cooling instability sets in very suddenly and shows some features of a radiative shock. Previous approximate treatments of the boundary layer^{12,13} included elements of the thermal instability but missed the importance of the advected entropy and the sudden cooling.

It is known that cataclysmic variables (CVs) radiate in hard X-rays, with spectra that appear to be consistent with a thermal bremsstrahlung spectrum with kT_{brems} of a few to tens of kiloelectronvolts. As the accretion rates of CVs decrease, and especially when they drop below $\dot{M} \lesssim 10^{-9.5} M_\odot \text{ yr}^{-1}$, their hard X-ray luminosity becomes a substantial fraction of the total flux³. Our results are consistent with these observations. The hot ($\sim 10^8$ K), optically thin boundary layers that we find for low values of $\dot{M}$ should produce abundant X-rays with $kT_{\text{brems}} \approx 10$ keV. The optically thick boundary layers at higher values of $\dot{M}$ will emit a far smaller fraction of their luminosity in X-rays. As ΔR_{BL} and T_{BL} both depend on $\dot{M}$, M_* , R_* , Ω_* and α , we should be able to compare our self-consistent solutions with observations and thus constrain the parameters of the accreting white dwarfs.

A possible difficulty in making such a comparison is that the hard X-ray emission is often variable (perhaps because of time-dependent features of the thermal instability), and thus cannot be fitted by our steady-state models. We must also caution that the radiative transfer we have used is somewhat simplistic. We have not, for instance, considered the scattering of horizontal to vertical flux (or vice versa) due to electron scattering, or energy transport through conduction. It is straightforward to include these effects, which will introduce modest changes, particularly in the thermal width of the boundary layer. We also intend to use a more detailed model of vertical radiative transfer

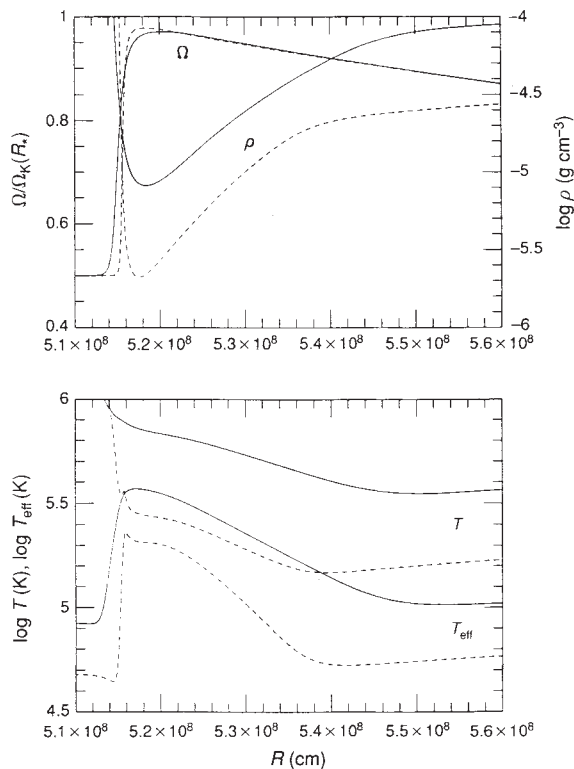


FIG. 1 Optically thick boundary layer region for two solutions with high mass accretion rates $\dot{M}$; the solid lines correspond to $\dot{M} = 10^{-7.5} M_\odot \text{ yr}^{-1}$, and the dashed lines to $\dot{M} = 10^{-8.5} M_\odot \text{ yr}^{-1}$. The accreting material flows from right to left, and meets the stellar surface at $R \approx 5 \times 10^8$ cm. The top panel shows the angular velocity of the accreting material Ω , and the density ρ . Ω is nearly keplerian in the disk, but drops abruptly to the stellar rotation rate $\Omega_* = 0.5\Omega_K(R_*)$ in the boundary layer. The bottom panel shows the central temperature T and the effective temperature T_{eff} . T rises gradually through the boundary layer and more rapidly as the material settles onto the star. T_{eff} peaks in the boundary layer, but is large over a fairly wide region, so that the boundary-layer luminosity is radiated from an area larger than that of the boundary layer itself.

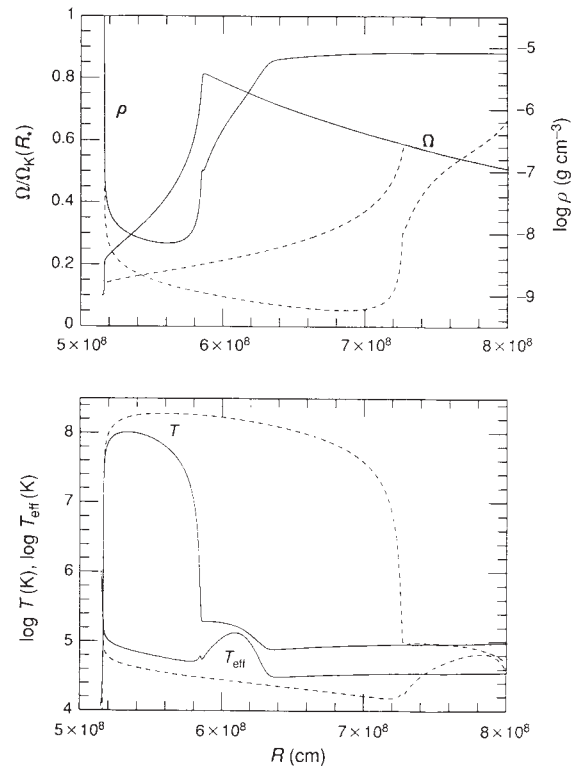


FIG. 2 Optically thin boundary layer region for two solutions with low mass accretion rates $\dot{M}$ and with a stellar rotation rate $\Omega_* = 0.1\Omega_K(R_*)$; the solid lines correspond to $\dot{M} = 10^{-9.5} M_\odot \text{ yr}^{-1}$, and the dashed lines to $\dot{M} = 10^{-10.5} M_\odot \text{ yr}^{-1}$. The boundary layers in these solutions are much wider than in the optically thick solutions shown in Fig. 1. The width increases rapidly with decreasing $\dot{M}$. The drop in density in the boundary layer is far more pronounced than in the optically thick solutions, and the central temperature jumps dramatically, exceeding 10^8 K before falling suddenly when the accreting material reaches the star. The effective temperature is low in the boundary layer because of the low optical depth, but is larger in a narrow spike at the inner edge of the boundary layer and in a small region just outside the boundary layer, where the optical depth is larger.

in order to calculate realistic continuum spectra. We will then be able to compare observed spectra and model predictions directly. We should be able to understand, for instance, why so few CVs have been detected in soft X-rays, as might be expected from an optically thick boundary layer.¹⁴

Although the models we have presented here are for CVs, we may be able to use the same techniques to construct models of other accretion disk systems. Accreting neutron stars and black holes are similar to CVs in many respects but differ by being radiation-dominated in the boundary-layer region^{1,2}. In analogy with the present results, we might expect these systems to have cooler, narrower boundary layers at high optical depth and hotter, wider boundary layers at low optical depth. There is evidence for this behaviour, with some galactic binary black-hole candidates^{15,16} as well as a few neutron star systems showing transitions from a soft spectrum at higher luminosities to a hard spectrum at lower luminosities.

A potentially important difference between neutron stars and black holes is in the nature of the inner boundary condition. Accretion onto an unmagnetized neutron star is similar to the white dwarf case considered here, with Ω changing from Ω_K to Ω_* within the boundary layer and v decelerating significantly when the material reaches the star. In contrast, a black hole has a 'soft' inner boundary, so that Ω is nearly equal to Ω_K down to the inner edge of the disk and then nearly constant to the horizon. Also, the radial velocity v increases monotonically inwards, passing through a sonic point¹⁰. (A neutron star can show similar behaviour if its radius is smaller than the radius of the last stable orbit, but this needs a soft equation of state and a slowly rotating star). This difference must lead to distinct signatures in the spectra, which we hope to discover using the techniques developed here.

Many models of active galactic nuclei postulate the presence of a supermassive black hole that creates a hot pair plasma¹⁷, but the mechanism by which the central engine maintains the required high temperatures has been unclear. By simply scaling from the white dwarf solutions presented here, we would expect optically thin boundary layers around black holes or neutron stars to produce temperatures as high as 10^{10} – 10^{11} K by steady accretion. This is hot enough for the creation of electron-positron pairs and could be relevant to pair production in active galactic nuclei. But our model does not include some effects that could be important at such high temperatures. For instance, the boundary layer would be cooled both by Compton scattering of soft radiation from elsewhere in the disk¹⁸, and by conduction. This will limit the maximum temperature attainable in these systems and should be included self-consistently in the models.

Finally, this work may also be applicable to accreting young stellar objects. Among this class of stars, the FU Orionis systems with high $\dot{M} \approx 10^4 M_\odot \text{ yr}^{-1}$ certainly have optically thick boundary layers, and T Tauri stars with $\dot{M} \lesssim 10^{-7} M_\odot \text{ yr}^{-1}$ may have optically thick or thin boundary layers^{19,20}. There is a good deal of observational data on the disks and boundary layers of these systems which could be exploited by computing self-consistent models. In a few T Tauri stars there is evidence that the accretion disk appears to terminate at a few stellar radii. We speculate that this may in some cases result from an optically thin boundary layer that happens to be several stellar radii wide. Such wide boundary layers are possible if $\alpha \geq 1$ and if the central star spins much below break-up as observed. $\square$

Received 15 January; accepted 18 March 1993.

- Shakura, N. I. & Sunyaev, R. A. *Astr. Astrophys.* **24**, 337–355 (1973).
- Pringle, J. E. *A. Rev. Astr. Astrophys.* **19**, 137–162 (1981).
- Patterson, J. & Raymond, J. C. *Astrophys. J.* **292**, 535–549 (1985).
- Popham, R. & Narayan, R. *Astrophys. J.* **370**, 604–614 (1991).
- Paczynski, B. *Astrophys. J.* **370**, 597–603 (1991).
- Narayan, R. *Astrophys. J.* **394**, 261–267 (1992).
- Popham, R. & Narayan, R. *Astrophys. J.* **394**, 255–260 (1992).
- Paczynski, B. & Bisnovatyi-Kogan, G. *Acta astr.* **31**, 283–291 (1981).
- Muchotrzeb, B. & Paczynski, B. *Acta astr.* **32**, 1–11 (1982).
- Abramowicz, M. A., Czerny, B., Lasota, J. P. & Szuszkiewicz, E. *Astrophys. J.* **332**, 646–658 (1988).

- Hubeny, I. *Astrophys. J.* **351**, 632–641 (1990).
- Tylenda, R. *Acta astr.* **31**, 267–281 (1981).
- Regev, O. *Astr. Astrophys.* **126**, 146–151 (1983).
- Patterson, J. & Raymond, J. C. *Astrophys. J.* **292**, 550–558 (1985).
- Sunyaev, R. A. & Trumper, J. *Nature* **279**, 506–508 (1979).
- McClintock, J. *Ann. N.Y. Acad. Sci.* **647**, 495–502 (1991).
- Rees, M. J. *A. Rev. Astr. Astrophys.* **22**, 471–506 (1984).
- Shapiro, S. L., Lightman, A. P. & Eardley, D. M. *Astrophys. J.* **204**, 187–199 (1976).
- Bertout, C., Basri, G. & Bouvier, J. *Astrophys. J.* **330**, 350–373 (1988).
- Hartigan, P. et al. *Astrophys. J.* **382**, 617–635 (1991).

ACKNOWLEDGEMENTS. We thank M. Abramowicz for useful comments. This work was supported in part by a grant from the U.S. NSF.

Shaping of diamond films by etching with molten rare-earth metals

S. Jin, J. E. Graebner, M. McCormack, T. H. Tiefel, A. Katz & W. C. Dautremont-Smith

AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

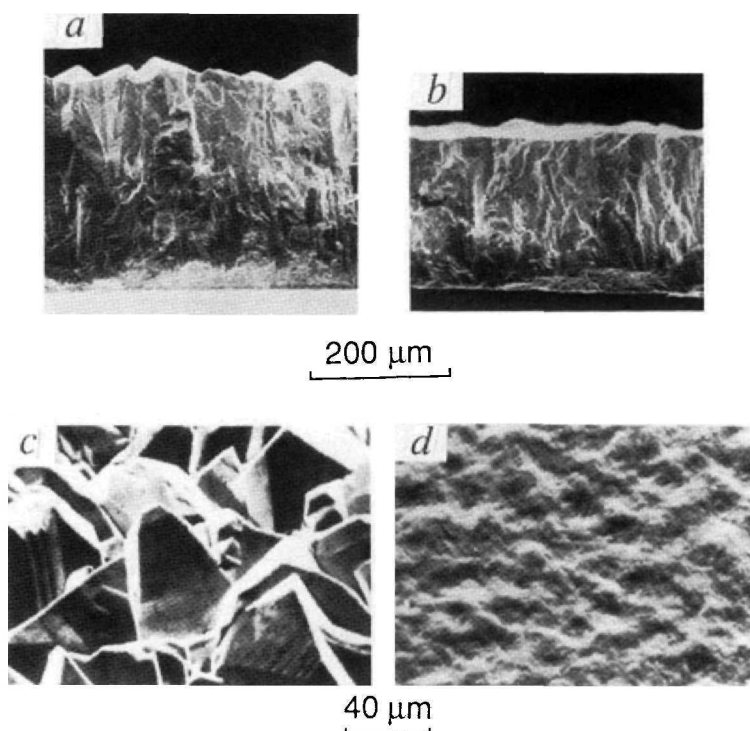
DIAMOND films prepared by chemical vapour deposition (CVD) have received considerable attention because of their high thermal conductivity, optical transparency, hardness, inertness and semiconducting properties^{1,2}. The control of film geometry (by thinning, polishing, patterning and shaping) is essential for most applications. Polishing by reaction with oxygen gas or ions usually results in pitting at grain boundaries^{3,4}, and other techniques^{5–8} are generally slow. Thinning or polishing of single-crystal diamond⁹ and CVD films^{4,10} by high-temperature diffusional reaction with solid transition metals has been reported previously. But these approaches require either a hydrogen atmosphere⁹, which is undesirable industrially, or high pressures (200–3,000 p.s.i.) during the heat treatment^{4,10}. We report here a method for shaping and thinning which does not suffer from these limitations. It involves the use of molten rare-earth metals such as cerium or lanthanum to etch away the diamond film surface. Substantial thinning and shaping can be achieved in just a few hours. We anticipate that this technique may open up new applications of diamond films in optics and biomedicine.

Free-standing CVD diamond films (1 cm × 0.5 cm × 250–300 μm thick; General Electric Superabrasives Company) were sandwiched between two sheets of $\sim 250\text{-}\mu\text{m}$ -thick rare-earth metal (cerium or lanthanum), and heat-treated under a light weight of $\sim 40\text{ g}$ (equivalent to a pressure of $\sim 1\text{ p.s.i.}$) at $\sim 920^\circ\text{C}$ for 4 h in an argon gas atmosphere.

After the reaction heat treatment, the residual (unreacted) and reacted rare-earth metals were removed from the diamond film surfaces by wet chemical etching using aqua regia ($\text{HNO}_3:\text{HCl}=1:3$). No detectable trace of Ce or La was found on the film surfaces after the chemical etching, either by energy-dispersive X-ray analysis (EDXA) in a scanning electron microscope or by Rutherford back-scattering analysis (F. A. Baiocchi, personal communication). The latter would have detected monolayer levels of Ce or La on diamond. The microstructure and the surface morphology of the films were investigated by scanning electron microscopy. The thermal conductivities of the films were measured both for the in-plane direction and the perpendicular (through-thickness) direction using techniques described previously^{2,11}.

An examination of available binary phase diagrams indicates that some molten rare-earth metals such as Ce (melting point 798°C) and La (melting point 918°C) can dissolve a very high percentage of carbon^{12,13} ($\sim 25\text{ at\%}$ at $\sim 920^\circ\text{C}$). This solubility is much larger than the solid solubility of carbon in transition metals such as iron or manganese ($\sim 6\text{--}12\text{ at\%}$ at 920°C). As the diffusivity of atoms in liquid metals is much higher than in solid metals, faster dissolution of carbon is also expected.

FIG. 1 Scanning electron micrographs of the CVD diamond films. Cross-sectional view before (a) and after (b) molten cerium etching; top view before (c) and after (d) molten lanthanum etching, showing the near-complete removal of rough growth facets.



From the cross-sectional microstructures of the fracture surfaces of the diamond film before and after treatment with molten cerium (Fig. 1a and b), it is evident that a substantial thinning (from $\sim 300 \mu\text{m}$ to $\sim 230 \mu\text{m}$ in thickness) has been accomplished. This thickness reduction, $\Delta t \approx 70 \mu\text{m}$, after only 4 hours at $\sim 920^\circ\text{C}$ is about 5 to 10 times faster than for diffusional reaction with solid metals (such as Fe, Mn) even with 2 to 3 orders of magnitude higher pressure applied during the solid metal processing.

Both the rough growth facets on the top surface and the fine-grained bottom region of the film have been removed by the molten rare earth processing, the mechanism of which is believed to be the diffusional transfer of carbon atoms in diamond to the molten metal. We have yet to determine whether the transfer of carbon goes through an intermediate stage of carbide formation. We are also investigating whether the rate of diamond removal is the same for both surfaces. The much larger grain boundary area in the bottom region of the CVD film may, for example, affect the reaction kinetics. The essentially complete removal of the rough growth facets by molten lan-

thanum processing (920°C for 4 h) is clearly illustrated in Fig. 1c and d. The top surfaces of the diamond film samples were scanned by a height-sensing stylus (DEKTA model No. 9000-50), and the results are shown in Fig. 2. Molten cerium etching yields similar results but with slightly rougher surface morphology than the lanthanum etching. Mechanical or laser polishing may be added, if necessary, to obtain a still smoother surface finish.

For fabrication of diamond-based semiconductor devices, thinning, polishing and patterning must be done at low temperature to minimize temperature- or diffusion-induced damage to doped regions, metallization and interconnections. We have achieved a marked reduction in thinning temperature (from about 900°C to $500\text{--}600^\circ\text{C}$) by using rare-earth/transition-metal alloys, such as 89% Ce/11% Ni (by weight) with a eutectic melting temperature of 477°C . The results (not shown here) are similar to those depicted in Fig. 1.

The reaction of CVD diamond with the molten rare-earth metals also reshaped the diamond films. Comparative micrographs of the samples (Fig. 3) show that the processing is

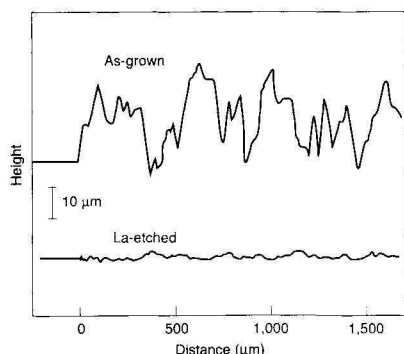


FIG. 2 Surface profile from the top surface of CVD diamond films.

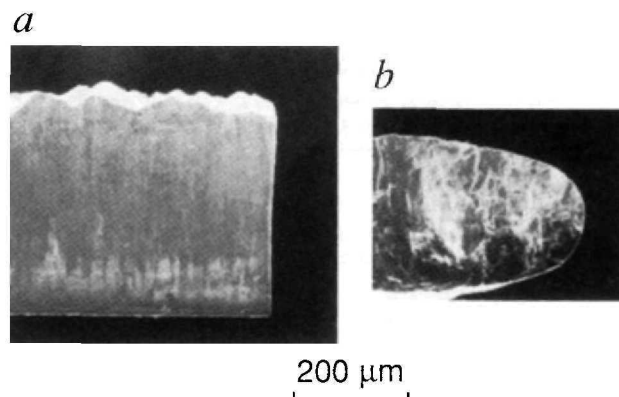


FIG. 3 Change in cross-sectional shape of the diamond film induced by molten Ce etching. a, Before reaction. b, After reaction.

effective not only in thinning and polishing diamond films, but also in eliminating all sharp corners and edges, and making them well rounded when excess molten cerium is provided beyond the edge of the film. The driving force for this effect as well as the afore-mentioned removal of growth-facets seems to be the reduction of surface energy in the diamond film during its reaction with the molten metal through the decrease of total surface area.

Because the atomic arrangement is less dense at grain boundaries and diffusion of atoms is usually faster there, metallic contamination at grain boundaries in the polycrystalline CVD diamond is a possibility during molten metal processing. If such contamination had occurred, increased phonon scattering at the grain boundaries would have resulted in a noticeable decrease in thermal conductivity. Measurements indicate that there is no degradation in either the through-thickness or in-plane thermal conductivity ($K_{\perp}$ and $K_{\parallel}$, respectively), implying that few additional phonon scattering centres have been introduced. In fact, etching the diamond films with molten cerium (top and bottom surfaces) considerably improved $K_{\perp}$, typically from ~ 14 to $\sim 20 \text{ W cm}^{-1} \text{ K}^{-1}$. (The highest room-temperature value obtained so far in the Ce-etched diamond films is $\sim 22 \text{ W cm}^{-1} \text{ K}^{-1}$, which is comparable to the thermal conductivity of the best type IIa single-crystal diamond.) The improvement is believed to be mainly due to the removal of the fine-grained, low-conductivity material from the bottom region². The in-plane thermal conductivity ($K_{\parallel}$, typically $\sim 13 \text{ W cm}^{-1} \text{ K}^{-1}$) in the Ce-etched diamond films is also improved, although conductivities are lower in this direction (as might be expected from the anisotropic microstructure), and the degree of $K_{\parallel}$ improvement by Ce-etching is considerably smaller than the $K_{\perp}$ improvement.

Our processing technique could be adopted for simultaneous processing of many diamond films, by interleaving them with thin sheets of rare-earth metals or immersing them in a bath of the molten metal. Shaping of diamond into non-flat geometries has not previously been investigated extensively. Our method might be applied to the fabrication of wires, needles, lenses and other shapes from CVD films. For complicated shapes, a solid-state reaction would be preferred, whereby the metal is pre-shaped and acts as a negative mould for the diamond body; but the rate of etching would be slower in this case. The large refractive index of diamond ($n=2.4$) makes it attractive for optical applications, as do the strength, high thermal conductivity and low thermal expansion coefficient. Pre-shaped rare-earth metal moulds with concave cavities might be used for shaping diamond films into an array of microlenses for optical device applications. □

Received 14 January; accepted 12 March 1993.

1. Angus, J. C. et al. *A. Rev. Mater. Sci.* **21**, 221-248 (1991).
2. Graebner, J. E. et al. *Nature* **359**, 401-403 (1992).
3. Harker, A. B. in *Applications of Diamond Films and Related Materials* (eds Tzeng, Y. et al.) 223-225 (Elsevier, New York, 1991).
4. Jin, S. et al. *Diamond Related Mater.* **1**, 949-953 (1992).
5. Hirata, A. et al. in *Applications of Diamond Films and Related Materials* (eds Tzeng, Y. et al.) 227-232 (Elsevier, New York, 1991).
6. Tokura, H. & Yoshikawa, M. in *Applications of Diamond Films and Related Materials* (eds Tzeng, Y. et al.) 241-248 (1991).
7. Yoshikawa, M. *Diamond Films Technol.* **1**, 1-29 (1991).
8. Silveri, P. *Indust. Diamond Rev.* **46**, 108-109 (1986).
9. Grigoriev, A. P. & Kovalsky, V. V. *Indiaqua* **39**, 47-54 (1984).
10. Jin, S. et al. *Appl. Phys. Lett.* **60**, 1948-1950 (1992).
11. Graebner, J. E. et al. *Appl. Phys. Lett.* **60**, 1576-1578 (1992).
12. *Binary Alloy Phase Diagrams*, 2nd Edn (eds Massalski, T. B.) 834 (ASM International, Metals Park, Ohio, 1990).
13. *Binary Alloy Phase Diagrams*, 2nd Edn (eds Massalski, T. B. et al.) 856 (ASM International, Metals Park, Ohio, 1990).

Prehistoric increases in the pH of acid-sensitive Swedish lakes caused by land-use changes

I. Renberg, T. Korsman & H. J. B. Birks*

Department of Ecological Botany, University of Umeå, S-901 87 Umeå, Sweden

* Botanical Institute, University of Bergen, N-5007 Bergen, Norway

IN the past few decades, there has been considerable debate about whether vegetational and soil changes associated with changing land-use can cause surface-water acidification in lakes^{1,2}. Although it is now widely accepted that the severe and extensive acidification that has occurred recently in southern Scandinavia, northern Britain and North America^{3,4} has been chiefly caused by atmospheric acid deposition, the role of changing land-use for moderate acidification is still not fully understood^{5,6}. Here we report analyses of sediment records from Swedish lakes, which provide evidence that land-use changes can have an important influence on the pH of acid-sensitive lakes. Following the expansion of an agrarian economy during the Iron Age (from about 2,500 years ago), pH increased from about 5.5 to about 6.5. Our results suggest that this pH increase was caused by burning, agriculture, forest grazing and other culture-related practices that increased the base saturation and pH of the soils, and enhanced the transport of base cations and nutrients from the soils to the surface waters, thus indicating these lakes may be naturally more acidic than had been thought. Following this period of alkalization, pH in many lakes fell to about 4.5 during the present century. Although cessation of former land-use practices could account for some of this change, the unprecedentedly low pH in recent years must be due predominantly to acid deposition.

In a previous palaeolimnological investigation⁷ of two lakes in southern Sweden (Lilla Öresjön and Lysevatten), an increase of diatom-inferred pH (DpH) from about 5.3 to 6.5 was recorded in sediment layers dated to $\sim 2,000$ years before present (BP). This increase was unexpected, as earlier studies from lakes in areas with poor shallow soils overlying granitic bedrock had shown gradually decreasing or rather stable pH over the last $>10,000$ years since deglaciation^{8,9}. Our investigation was designed to assess whether this unexpected pH increase was general for a larger area in southern Sweden and to test the hypothesis that the change was caused by human land-use and disturbance of the natural forests and soils.

Sediment cores were taken from 12 additional lakes (Fig. 1) for analysis of fossil diatoms as a tool for reconstructing lake pH history¹⁰. Because of the focus of this study, selected core sections were analysed and only in four lakes has the whole postglacial period (Holocene) been analysed.

The lakes studied are typical for southwest Sweden, a region that suffers heavily from surface-water acidification. In the study area, 50% of the lakes ($\sim 3,000$ lakes) are classified as acidified by the Swedish Environmental Protection Agency¹¹. Lakes are defined as acidified when the average alkalinity has decreased by $>25\%$ since pre-industrial times, corresponding to a pH decline of >0.125 pH units¹¹. The catchments of the studied lakes have thin till soils and at present support forests predominantly of *Pinus sylvestris*, *Picea abies* and *Betula* spp. There are no peatlands near the lakes.

In 12 of the 14 lakes, a major change in the diatom flora and a significant increase in DpH occurred at depths between 40 and 100 cm in the sediments (Fig. 2). The diatom flora, which had been dominated by acidophilous, benthic species, such as *Frustulia rhomboides* agg., *Brachysira brebissonii*, *B. vitrea*, *Aulacoseira distans* var. *navalis*, *A. perglabra*, *A. lirata* and *Eunotia* spp., shifted to a more alkaliphilous flora rich in plankton, with taxa such as *Cyclotella* spp., *Asterionella formosa* and,

in some lakes, *Aulacoseira ambigua*. DpH increased by 0.5–1.4 pH units, from 5.2–5.6 to 6.0–6.8 (Fig. 2). In two lakes there was no significant change; the pH in these lakes has always been above 6 in the past (Fig. 2). The pH increase was ^{14}C -dated in six lakes to 2,300–1000 BP. As the changes in different lakes were not synchronous, climate control is unlikely.

Statistical analyses of pollen, charcoal, loss-on-ignition (LOI) and DpH on data from the same sediment cores from six of the lakes strongly support the hypothesis that this pH increase was caused by changes in land-use (Table 1, Fig. 3), with 66–95% of the variance in DpH at a site being accounted for by explanatory variables such as *Juniperus*, *Alnus*, Apophyte or *Hordeum* pollen values, *Polypodiaceae* and *Lycopodium annotinum* spore frequencies, and LOI. In many instances a single explanatory variable captures 50–79% of the DpH variance at the sites (Table 1). Because of multicollinearity between many of the explanatory variables, the multiple regression analyses of Table 1 cannot distinguish between variables that are nearly as good predictors for DpH as the ones selected by forward selection (Table 1). For example, at Stora Skärsjön Poaceae is almost as good as Apophytes as an explanatory variable, at Lilla Holmevatten *Picea* is almost as good as *Alnus*, at Stora Langen Poaceae is almost as good as LOI, and at Iglasjön Ericaceae, *Calluna vulgaris*, and *Sphagnum* are all almost as important as Polypodiaceae, whereas charcoal is almost as good as *Lycopodium annotinum* as a predictor of DpH.

Despite these problems of multicollinearity, there is a variety of strong stratigraphical signals for increased land-use and disturbed vegetation simultaneous with the pH increase: recession of the natural forest (indicated by Poaceae), expansion of shrub vegetation (*Juniperus communis*), increased frequency of cereals and weeds (*Hordeum*, *Rumex acetosa acetosella*), increased concentrations of charcoal, and, in some lakes, decreased LOI values indicating soil erosion (Table 1, Fig. 3). We interpret these changes as the effects of an expanding agrarian culture during the Iron Age. This culture was largely based on animal husbandry, and burning was important for land clearance to create and maintain fodder-producing land. Burning became much more common than in the previous natural deciduous forest where the natural fire frequency was low, as indicated by the charcoal analyses.

It is likely that fire was the principal mechanism for the pH increase we have recorded. Burning destroys acid raw humus, and releases base cations and nutrients bound in biomass and organic soil layers, leading to increased base saturation of soils

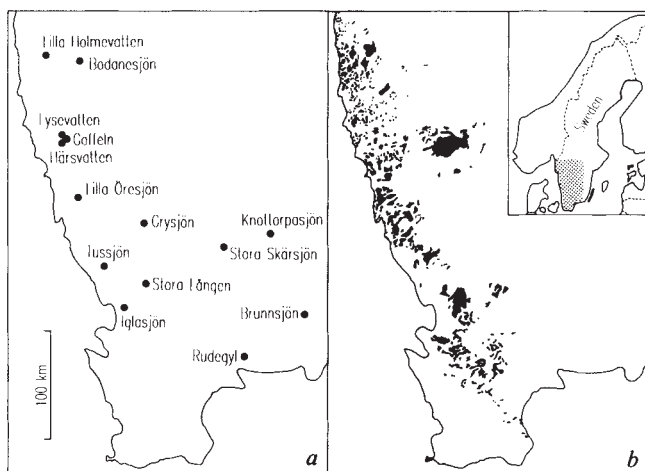


FIG. 1 Map of South Sweden with (a) studied lakes and (b) the extent of *Calluna* heaths according to a map²³ from 1897. At that time, the heathland area had already begun to decrease. Today all former heaths are covered by forests. All the studied lakes are small (<0.5 km²) headwater lakes, and all except Knottorpasjön have clear water today.

TABLE 1 Statistical analysis

Site	Explanatory variable	Percentage variance explained ($R^2 \times 100$)	p^*	r	Regression coefficient†
Härsvatten	<i>Juniperus</i>	74	0.002	0.86	0.86
Lilla Holmevatten	<i>Alnus</i>	69	0.002	-0.83	-0.99
	LOI	9	0.002	0.11	-0.35
	Total model	78	0.004	0.88	—
Stora Skärsjön	Apophytes‡	79	0.002	0.89	0.62
	<i>Alnus</i>	9	0.05	0.34	0.39
	<i>Hordeum</i>	6	0.002	0.72	0.37
	Total model	95	0.002	0.97	—
Iglasjön	Polypodiaceae	50	0.002	0.71	0.80
	<i>Lycopodium annotinum</i>	16	0.024	-0.24	-0.41
	Total model	66	0.002	0.81	—
Stora Langen	LOI	60	0.002	-0.78	-0.84
Rudegyl	Apophytes‡	73	0.002	0.85	0.66
	<i>Corylus</i> type	19	0.002	-0.22	-0.23
	Total model	92	0.002	0.96	—

Results of forward selection^{21,22} within constrained multiple regression analysis²¹ to find a minimal set of statistically significant explanatory predictor variables among the major terrestrial pollen and spore types, LOI and charcoal for the observed stratigraphical changes in DpH with the constraint of sample depth. DpH based on weighted averaging regression and calibration¹⁰ of the diatom data is the sole response variable. Statistical significance of each explanatory variable was assessed by means of a restricted Monte Carlo permutation test for time-ordered stratigraphical data²². We used 499 permutations and any explanatory variable with an exact Monte Carlo significance value of $p \leq 0.05$ was considered significant. The significance of the overall multiple regression model for sites where there are two or more significant explanatory variables was also assessed by restricted Monte Carlo permutation tests (499 permutations).

* Exact Monte Carlo significance level²¹.

† Regression coefficients of the regression of standardized response variable (DpH) on standardized explanatory variables.

‡ Mainly Poaceae, *Rumex acetosa/acetosella*, *Plantago media/major* and *Urtica* type.

and increased soil pH^{12,13}. Through leaching of the ash, bases and nutrients are also lost from the soil system to the surface waters, where pH and alkalinity may increase. During fires large amounts of particulates are emitted to the air, and in a landscape with abundant fires, atmospheric deposition of nutrients and bases may also benefit unburnt areas^{14,15}. Some lakes may also have been affected by direct run-off from cultivated land, although most lakes have rocky catchments with thin soils that are unsuitable for agriculture. Grazing and trampling by cattle may also have increased the decomposition of vegetation and organic soil fractions, and prevented the accumulation of acid humus soils.

Land-use expanded and became progressively more intense with time, resulting in large areas with a sparse forest cover or heathland in south Sweden. Several, but not all, of the studied lakes are situated in areas that were *Calluna vulgaris*-dominated heaths until the late nineteenth century (Fig. 1). It is possible that these heathlands began to be formed about 2,000 BP. According to historical sources¹⁶, the heaths were burnt every 5–10 years, thereby resulting in recurrent regeneration of soil pH, higher pH of run-off water, and a constant supply of base cations and nutrients to the lakes. Maximum deforestation was reached in the middle of the nineteenth century, after which reforestation programmes commenced¹⁶.

The cultural alkalization period we have described succeeded nearly 10,000 years of long-term natural acidification from about pH 7 to 5.5 (Fig. 2) caused by progressive leaching and loss of base cations from the soils, leading to soil acidification and dilution of run-off. The alkalization period continued until this century, when pH started to decrease.

These results raise again the much discussed question of whether changes in land-use, vegetation and soils associated with the shift from an agrarian to an industrial economy during this century could have contributed to recent surface-water acidification^{1,2,17,18}. As these landscape changes have occurred simultaneously with the increased atmospheric deposition of

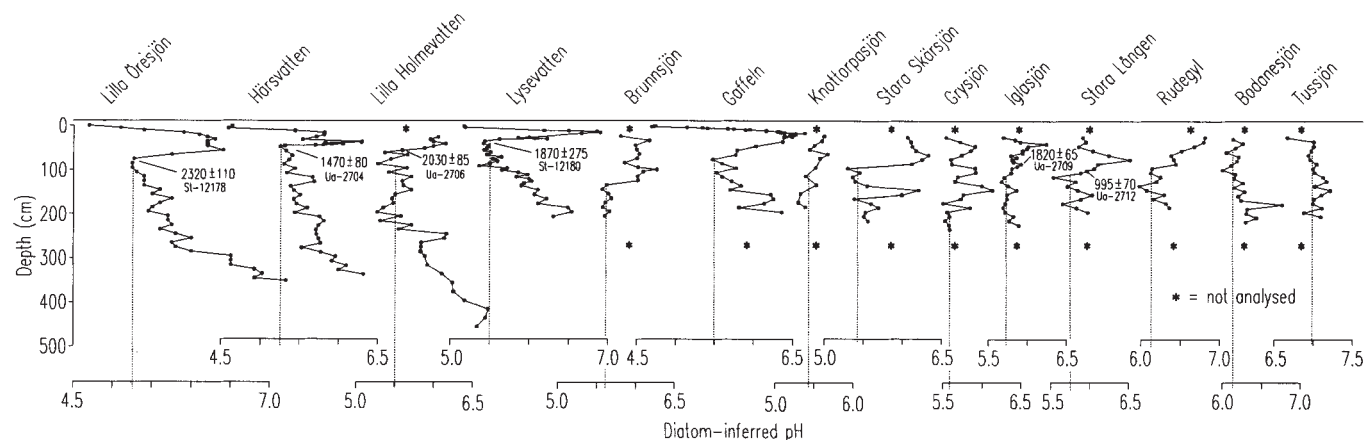


FIG. 2 Diatom-inferred past pH (DpH) plotted against depth in sediment cores from lakes in southern Sweden. DpH was calculated using weighted averaging regression and calibration and the SWAP calibration data set from Britain, Norway and Sweden^{10,24}. Recent sediments were sampled with a freeze corer and older sediments with a Russian peat corer from lake-ice. Dates are radiocarbon years BP. Preliminary reconstructions using a new

calibration dataset from northern Sweden (T.K., unpublished work) give slightly higher DpH values (~0.3 pH) than the DpH values presented here. The stratigraphical trends in DpH are identical, however, using either calibration set. The aim was to study the prehistoric pH increase and its causes. Therefore, the upper and lower core sections of most lakes were not analysed.

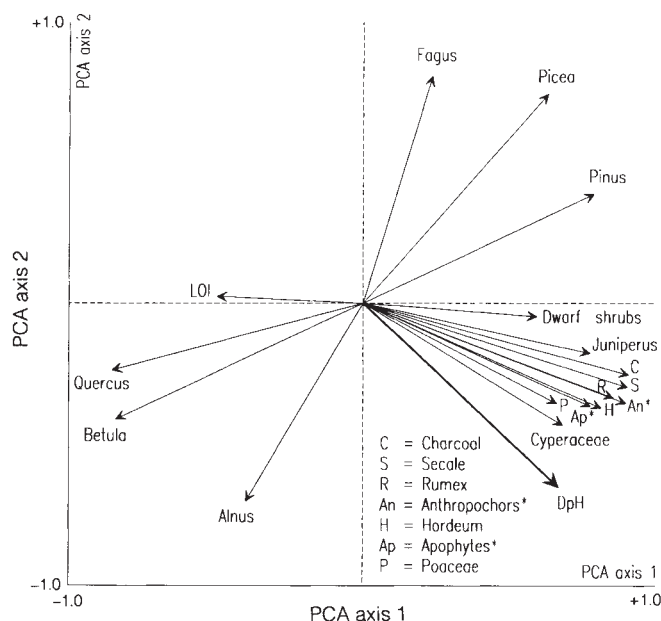


FIG. 3 Correlation principal components biplot²² of the explanatory variables or predictors (major terrestrial pollen and spore types, LOI, charcoal) of DpH at Stora Skärsjön (PCA, principal component analysis). DpH and the pollen groups Apophytes (native plants favoured by anthropogenic activity, such as *Poaceae*, *Rumex acetosa/acetosella*) and Anthrophochors (crop plants and alien weeds) were included as 'passive' variables in the analysis. A passive variable has no influence on the analysis but is added to the biplot afterwards by the use of transition formulae²¹. Apophytes and Anthrophochors are marked with an asterisk. In this biplot, variables with arrows that make an acute angle are positively correlated, with a higher correlation the longer the arrows are. Variables with obtuse-angled arrows have negative correlations. The biplot is based on a principal component analysis of a correlation matrix of the 15 active variables. Axis 1 represents 52.1% of the total variance whereas axis 2 captures 18.9%. Subsequent axes are very small (axis 3, 11.7%; axis 4, 5%) and are not considered here. There are high positive correlations between DpH, *Poaceae*, Apophytes, *Hordeum*, Anthrophochors, *Secale*, charcoal, *Rumex*, *Juniperus communis* and dwarf shrubs, high negative correlations between DpH and *Quercus*, *Betula* and LOI, and low correlations between DpH and *Alnus* and between DpH, *Picea* and *Fagus*.

strong acids, this study does not permit a critical evaluation of this question. However, a drop in surface-water alkalinity and pH to the pre-land-use level of about 5.5 might be expected, as burning and natural fires are extremely rare today and the cultural landscape of heathland, meadows and pastures is largely replaced by coniferous forests.

Although our work shows that land-use is important in influencing lake-water pH, it does not suggest that the severe recent acidification with pH values of about 4.5 and fish kills in numerous southern Swedish lakes are caused by land-use changes. Such low pH values never existed in these lakes before (Fig. 2), and this severe pH decline is only found in areas with high acid deposition. Moreover, there is overwhelming evidence from a series of critically designed studies^{8,19,20} that acid deposition is the main cause of recent surface-water acidification. The results presented here are relevant to discussions about liming policy and conservation of cultural landscapes in acid-sensitive regions. Our data raise the question of whether it is possible to recreate the chemical conditions in lakes that resulted from cultural practices no longer existing today. In practice, lakes cannot be separated from their catchments in such management and recreation policies. □

Received 10 February; accepted 18 March 1993.

- Rosenqvist, I. T. *Sci. tot. Envir.* **10**, 39–49 (1978).
- Krug, E. C. & Frink, C. R. *Science* **221**, 520–525 (1983).
- Mason, B. J. (ed.) *The Surface Waters Acidification Programme* (Cambridge Univ. Press, 1990).
- National Acid Precipitation Assessment Program *Acidic Deposition: State of Science and Technology, Vol. 2: Aquatic Processes and Effects* (NAPAP, Washington DC, 1991).
- Krug, E. C. *J. Hydrol.* **128**, 1–27 (1991).
- Battarbee, R. W. *Quat. Proc.* **2**, 101–110 (1992).
- Renberg, I. *Phil. Trans. R. Soc. Lond.* **B327**, 357–361 (1990).
- Charles, D. F. et al. in *Acid Precipitation, Vol. 4: Soils, Aquatic Processes, and Lake Acidification* (eds Norton, S. A., Lindberg, S. E. & Page, A. L.) 207–276 (Springer, New York, 1989).
- Battarbee, R. W. *Phil. Trans. R. Soc. Lond.* **B327**, 339–347 (1990).
- Birks, H. J. B. et al. *Phil. Trans. R. Soc. Lond.* **B327**, 263–278 (1990).
- Monitor 12. *Acidification and Liming of Swedish Freshwaters* (Swedish Environmental Protection Agency, Solna, 1991).
- Viro, P. J. in *Fire and Ecosystems* (eds Kozlowski, T. T. & Ahlgren, C. E.) 7–72 (Academic, New York, 1974).
- MacLean, D. A. et al. in *The Role of Fire in Northern Circumpolar Ecosystems* (eds Wein, R. W. & MacLean, D. A.) 111–132 (Wiley, Chichester, 1983).
- Chapman, S. B. *J. Ecol.* **55**, 677–689 (1967).
- Gimingham, C. H. *Ecology of Heathlands* (Chapman & Hall, London, 1972).
- Malmström, C. *Medd. Statens Skogsförsöksanstalt* **31**, 171–300 (1939).
- Overreim, L. N., Seip, H. M. & Tøllan, A. *Acid Precipitation—Effects on Forest and Fish SNSF-Project Rep. 1972–1980* (Norwegian Institute for Water Research, Oslo, 1980).
- Norton, S. A., Brakke, D. F. & Henriksen, A. *Sci. tot. Envir.* **83**, 113–125 (1989).
- Battarbee, R. W. et al. *Palaeolimnology and Lake Acidification* (Royal Society, London, 1990).
- Birks, H. J. B. et al. *J. Paleolimnol.* **4**, 69–85 (1990).
- ter Braak, C. J. F. *CANOCO Version 3.1* (Microcomputer Power, Ithaca, New York, 1990).

22. ter Braak, C. J. F. *Update Notes: CANOCO Version 3.10* (Agricultural Mathematics Group, Wageningen, 1990).
 23. Schager, N. Ymer **29**, 309–335 (1909).
 24. Stevenson, A. C. et al. *The Surface Waters Acidification Project: Modern Diatom/Lake-Water Chemistry Data-Set* (Ensis, London, 1991).

ACKNOWLEDGEMENTS. We thank J.-E. Wallin for help with the pollen analyses, A.-B. Lindström for charcoal analyses, and U. Segerström and D. Lindberg for fieldwork assistance. This work was supported by the Swedish Environmental Protection Agency, the Swedish Natural Science Research Council, Centre for Environmental Research, University of Umeå, and the Surface Waters Acidification Programme.

A 6,000-year sedimentary molecular record of chemocline excursions in the Black Sea

Jaap S. Sinninghe Damsté*†, Stuart G. Wakeham‡, Math E.L. Kohnen†§, J.M. Hayes|| & Jan W. de Leeuw*†

* Netherlands Institute of Sea Research (NIOZ), Division of Marine Biogeochemistry, PO Box 59, 1790 AB Den Burg, Texel, The Netherlands

† Organic Geochemistry Unit, Delft University of Technology, de Vries van Heystplantsoen 2, 2628 RZ Delft, The Netherlands

‡ Skidaway Institute of Oceanography, PO Box 13687, Savannah, Georgia 31416, USA

|| Biogeochemical Laboratories, Departments of Geological Sciences and of Chemistry, Indiana University, Bloomington, Indiana 47405, USA

THE Black Sea is the world's largest anoxic basin; it is also a contemporary analogue of the environment in which carbonaceous shales and petroleum source beds formed¹. Recently, Repeta et al.^{2,3} reported that anoxygenic photosynthesis may be an important component of carbon cycling in the present Black Sea, owing to a shoaling of the chemocline and consequent penetration of the photic zone by anaerobic waters in the past few decades^{4,5}. It has been suggested⁴ that this was due to an anthropogenic decrease in freshwater input to the Black Sea, although natural causes were not ruled out. Here we report the distributions of sequestered photosynthetic pigments in eight core samples of sediments from the Black Sea ranging in age from zero to 6,200 years before the present. Our results show that photosynthetic green sulphur bacteria (Chlorobiaceae) have been active in the Black Sea for substantial periods of time in the past. This finding indicates that the penetration of the photic zone by anaerobic waters is not a recent phenomenon, and suggests that natural causes for shoaling of the chemocline are more likely than anthropogenic ones⁴.

The Pleistocene and Holocene sedimentary record of the Black Sea^{6,7} reveals that the basin changed from a freshwater lake to the present anoxic marine basin ~9,000 years before present (BP), when the post-glacial sea-level rise reconnected the Black Sea with the Mediterranean. The base of a microlaminated organic-rich sapropel layer (unit 2) marks the development of anoxic bottom waters ~7,000 years BP in the centre of the basin. On the slope of the basin, deposition of unit 2 started ~1,000 years later, indicating a slow rise of the chemocline⁷ and thus increasing depletion of oxygen. It is generally assumed that this rise continued during deposition of unit 1 (ref. 8), a finely laminated coccolith marl formed when increasing salinity allowed the invasion of surface waters by the marine alga, *Emiliani huxleyi*^{9,10}. The present-day chemocline (as defined by the depth at which hydrogen sulphide first appears and oxygen disappears in the water column) is at 80–100 m (refs 2, 4, 5, 11). Since 1969 it has risen by 20–50 m within the water column, and it has been hypothesized⁴ that this change is due to

TABLE 1 ¹³C contents of selected hydrocarbons released by desulphurization of macromolecular aggregates

Age (kyr)	<i>n</i> -C ₂₂ – <i>n</i> -C ₃₄ (%)	Phytane (%)	$\delta^{13}\text{C}^*$		
			β,β -carotane (%)	isorenieratene (%)	Octadecahydro isorenieratene (%)
BC10 (0–2.5 cm)	0–0.1	–29.7––32.2	–29.9	–29.7	–16.1
BC10 (15–20 cm)	1.0–1.2	–29.0––30.8	n.d.	–30.2	–14.8
GCC20 (17–20 cm)	5.2–6.2	–29.3––30.6	–29.5	–25.7	–17.4

* $\delta^{13}\text{C} = [(R_x - R_s)/R_s]$ where $R = {}^{13}\text{C}/{}^{12}\text{C}$, x designates sample, s designates PDB standard and $R_s = 0.0112372$. They were measured using conditions reported²⁹.

anthropogenic effects, specifically a reduction of inputs of fresh water. There are, however, also reports^{12,13} that the depth of the chemocline and pycnocline has oscillated by as much as 80 m on a timescale of months, suggesting that natural factors must also be important.

The location of the chemocline within the photic zone of the present-day Black Sea supports the world's largest and deepest population of anoxygenic photosynthetic bacteria^{2,3}. Detailed analysis of photosynthetic pigments in samples taken in and below the suboxic zone overlying the sulphide-containing anoxic deeper water revealed the presence of compounds (including the di-aromatic carotenoid isorenieratene I, Fig. 1) characteristic of the obligate photoautotrophs *Chlorobium phaeobacterioides* and *C. phaeovibrioides*. The bacteria were recently isolated and identified as low-light-adapted strains of *C. phaeobacterioides*¹⁴. These bacteria can thrive in water with exceptionally low light intensities, such as those at the chemocline in the Black Sea, because they have enhanced concentrations of light-harvesting pigments, specifically isorenieratene and β -isorenieratene in combination with bacteriochlorophyll-*a*¹⁴. It has been suggested that the presence of a zone of anoxygenic photosynthesis is a recent development and results from the shoaling of the chemocline over recent decades^{2,3,5}.

To test this hypothesis, we examined remnants of aromatic and other carotenoids in sediments from the Black Sea recovered during the 1988 expedition of RV *Knorr* to the Black Sea. As shown in Fig. 2, stratigraphic positions of the samples range from the uppermost unit 1 to well into unit 2. In anoxic sediments, functionalized lipids may react with reduced inorganic sulphur species to form low-molecular-weight organic sulphur compounds¹⁵ or sulphur-rich macromolecular aggregates^{16,17} depending on the number and position(s) of the functional group(s) in the molecule. Because carotenoids possess multiple double bonds, they preferentially accumulate in macromolecular fractions¹⁷. The presence of sulphur compounds, probably derived from natural sulphurization of lipids of diatoms, in both unit 1 and unit 2 sediments (ref. 18; and S.G.W., M.E.L.K. and J.S.D., manuscript in preparation) demonstrates that this process is active in Recent Black Sea sediments.

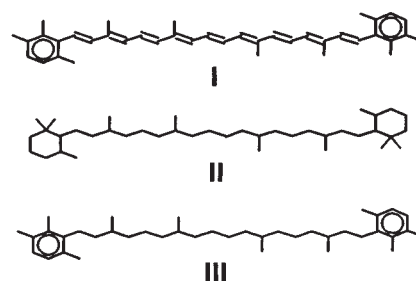


FIG. 1 Structures of compounds cited in the text. I, isorenieratene; II, β,β -carotane, III, octadecahydro isorenieratene.

§ Present address: Koninklijke/Shell Exploratie en Productie Laboratorium, PO Box 60, 2280 AB Rijswijk, The Netherlands.

Desulphurization of sulphur-rich macromolecular aggregates present in the polar fraction of the sediment extracts of our samples yielded, in addition to a series of linear and isoprenoid alkanes, the (partially) hydrogenated carotenoids, β,β -carotane (II) and, in most cases, octadecahydro-isorenieratene (III). The molecular structure of this latter component is highly specific and indicates the presence of photosynthetic green sulphur bacteria in the palaeoenvironment. The consistent enrichment of ^{13}C in this di-aromatic carotenoid relative to algal-derived sulphur-bound n -alkanes, phytane and β,β -carotane throughout the analysed section of the sediment (Table 1) confirms its origin from green sulphur bacteria. If these organisms grow as autotrophs, they fix carbon by means of the reverse tricarboxylic acid cycle leading to biomass anomalously enriched in ^{13}C (refs 19, 20). A similar level of ^{13}C -enrichment has been reported in sedimentary di-aromatic carotenoids¹⁷ and their presumed degradation products²¹. Notably, the green photosynthetic bacterial lipids, with a $\delta^{13}\text{C}$ value of about -16% relative to the PDB standard, are depleted in ^{13}C by only 3.5% relative to dissolved CO_2 ($\delta^{13}\text{C} = -12.5\%$) at 80 m depth (temperature $T = 8^\circ\text{C}$) in the present water column^{22,23}. Relative to bicarbonate ($\delta^{13}\text{C} = -1.5\%$), the isotopic fractionation is 14.5%. This observation provides the best information at present available about the isotopic fractionation of the reverse tricarboxylic acid cycle in natural marine environments.

The molecular and isotopic evidence thus indicates that anaerobic conditions have prevailed at some point within the

photic zone for substantial periods in the past 6,200 years, and that the present situation, in which shoaling of the chemocline has allowed photosynthetic green sulphur bacteria to thrive as autotrophs^{2,3,14}, is not unique. This finding is intriguing because it has been generally assumed that the rise of the chemocline occurred recently, but is supported by recent evidence for a relatively stable oxic-anoxic boundary in the Black Sea²⁴ and the isolation of intact isorenieratene (I) from Black Sea sediments (D. J. Repeta, personal communication). It follows that the postulated gradual ascent of the chemocline through time as determined by mass balance calculations⁸ should be reconsidered, and that anthropogenic causes for shoaling of the chemocline⁴ are not likely. We do not suggest, however, that the lower part of the photic zone has been permanently anoxic. The molecular-isotopic evidence records the common but not necessarily constant presence of Chlorobiaceae; frequent, rapid fluctuations in the depth of the chemocline could produce the same effect.

Our data also reveal the absence of an oxic-anoxic interface in the photic zone during deposition of the unit 1 transition sapropel and unit 2A sediments. The evidence is that the concentration of the di-aromatic carotenoid drops to zero whereas that of sulphur-bound β,β -carotane (derived from algae living in the upper photic zone) remains unaffected (Fig. 2). This is in partial agreement with the recent proposition that Black Sea water was oxic during deposition of unit 2 (ref. 25). It has been suggested that the prolific growth of *E. huxleyi* during the past 3,000 years was triggered by increased salinities of the photic zone^{9,10}. However, our di-aromatic carotenoid record suggests that the structure of the water column may also have been important (Fig. 2). If the chemocline rises to a level where winter mixing processes redistribute nutrients from the chemocline into the photic zone, the coccolithophorid blooms, which are characteristic of unit 1, might result²⁶. A consequence of the erosion of the chemocline would be mass mortality of the strictly anaerobic Chlorobiaceae. Indeed, the dark laminae within the coccolith oozes of unit 1 have been interpreted as products of seasonal mass mortality of photosynthetic bacteria²⁷, although an alternative explanation is that they are associated with the spring diatom bloom²⁸. If decreased mixing of the upper water column occasionally led to the absence of either of these events, or if Chlorobiaceae were occasionally absent because of a relatively deep chemocline, this would inhibit the formation of dark lamina. Possibly this is the cause of the discrepancy between the results of radiocarbon and varve chronology¹⁰ and the incomplete annual record in sediment traps²⁸.

The increase in concentration of octadecahydroisorenieratene in unit 2B indicates anoxia in the photic zone during deposition of these sediments. Although this situation is comparable with that existing today, any pumping of nutrients into the photic zone did not lead to blooms of *E. huxleyi*, because the salinity of the photic zone was too low. Replacement of fairly fresh bottom waters by incursion of denser Mediterranean waters into the deep basin as the Black Sea became increasingly brackish could have brought the chemocline into the photic zone.

The occurrence of anoxygenic photosynthesis during substantial periods over the past 6,200 years BP suggests that anaerobic photosynthesis may have been important in carbon cycling in the Black Sea. $\square$

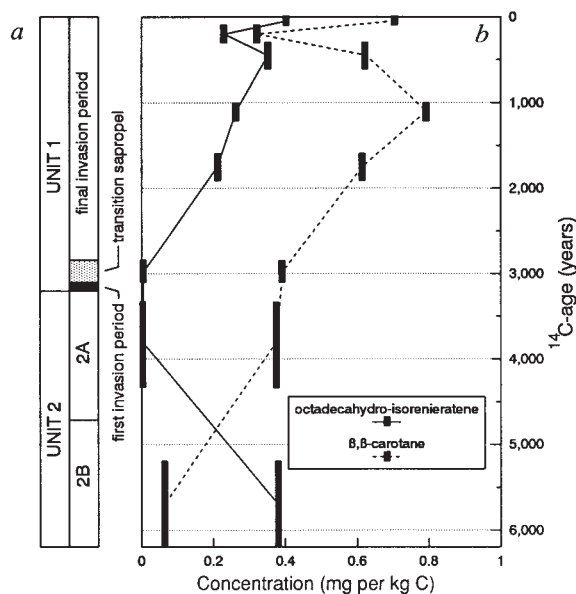


FIG. 2 a, Lithology of Holocene Black Sea sediments¹⁰. b, Concentrations of octadecahydro-isorenieratene (III) and β,β -carotane (II) released by Raney nickel desulphurization of macromolecular aggregates present in the polar fractions of the extracts. Samples were obtained from box cores and a giant gravity core taken at the abyssal plain (43°N , $32-34^\circ\text{E}$) at water depths greater than 2,000 m. Ages are estimated on basis of ^{14}C radio-chronology^{30,31}. Sizes of the bars indicate the time span covered by the sediment sample analysed. Polar fractions were separated from extracts (obtained by sonication with mixtures of CH_3OH and CH_2Cl_2) by column chromatography and desulphurized with Raney nickel¹⁷. Released hydrocarbons were separated by column chromatography, analysed by gas chromatography (GC) and by GC-mass spectrometry, and quantified by addition of the internal standard, 5-(1,1-d₂-hexadecyl)-2,3-dimethylthiophene before desulphurization. GC conditions: Carlo Erba 5300 instrument fitted with a 25 m $\times$ 0.32 mm CP-Sil 5 (film thickness 0.12 μm) fused-silica capillary column, temperature programmed from 70°C to 130°C at $10^\circ\text{C min}^{-1}$ and then to 320°C (hold time 20 min) at 4°C min^{-1} . Mass spectrometry conditions: VG-70S instrument, ionization voltage 70 eV, mass range m/z 40–800, cycle time 1.8 s.

Received 8 January 1992; accepted 6 April 1993.

1. Demaison, G. J. & Moore, G. T. *Bull. Am. Assoc. Petrol. Geol.* **64**, 1179–1209 (1980).
2. Repeta, D. J., Simpson, D. J., Jørgensen, B. B. & Jannasch, H. W. *Nature* **342**, 69–72 (1989).
3. Repeta, D. J. & Simpson, D. J. *Deep Sea Res. (Suppl. 2)* **38**, S969–S984 (1991).
4. Murray, J. W. et al. *Nature* **338**, 411–413 (1989).
5. Jørgensen, B. B., Fossing, H., Wirsen, C. O. & Jannasch, H. W. *Deep Sea Res. (Suppl. 2)* **38**, S1083–S1103 (1991).
6. Ross, D. A., Degens, E. T. & McIlvaine, J. *Science* **170**, 163–165 (1970).
7. Ross, D. A. & Degens, E. T. in *The Black Sea-Geology, Chemistry and Biology* (eds Degens, E. & Ross, D.) 183–199 (Mem. No 20, Am. Ass. Petrol. Geol., Tulsa, Oklahoma, 1974).
8. Deuser, W. G. in *The Black Sea-Geology, Chemistry and Biology* (eds Degens, E. & Ross, D.) 133–136 (Mem. No 20, Am. Ass. Petrol. Geol., Tulsa, Oklahoma, 1974).

9. Bukry, D. in *The Black Sea-Geology, Chemistry and Biology* (eds Degens, E. & Ross, D.) 253–263 (Mem. No 20, Am. Ass. Petrol. Geol., Tulsa, Oklahoma, 1974).
10. Hay, B. J., Arthur, M. A., Dean, W. E., Neff, E. D. & Honjo, S. *Deep Sea Res. (Suppl. 2)* **38**, S1211–S1235 (1991).
11. Tugrul, S., Basturk, O., Saydam, C. & Yilmaz, A. *Nature* **359**, 137–139 (1992).
12. Bryantsev, D. A., Fashchuk, D. Ya., Ayzatullin, A., Bagotskiy, S. V. & Leonov, A. V. *Oceanology* **28**, 180–185 (1988).
13. Kempe, S., Liebezeit, G., Diercks, A.-R. & Asper, V. *Nature* **346**, 419 (1990).
14. Overmann, J., Cypionka, H. & Pfennig, N. *Limnol. Oceanogr.* **37**, 150–155 (1992).
15. Brassell, S. C., Lewis, C. A., de Leeuw, J. W., de Lange, F. & Sinninghe Damsté, J. S. *Nature* **320**, 160–162 (1986).
16. Kohnen, M. E. L., Sinninghe Damsté, J. S. & de Leeuw, J. W. *Nature* **349**, 775–778 (1991).
17. Kohnen, M. E. L. *et al. Science* **256**, 358–362 (1992).
18. Kohnen, M. E. L. *et al. Geochim. cosmochim. Acta* **54**, 3053–3063 (1990).
19. Quandt, I., Gottschalk, G., Ziegler, H. & Stiehler, W. *FEMS Microbiol. Lett.* **1**, 125–128 (1977).
20. Sirevag, R., Buchanan, B. B., Berry, J. A. & Troughton, J. H. *Arch. Microbiol.* **112**, 35–38 (1977).
21. Summons, R. E. & Powell, T. G. *Nature* **319**, 763–765 (1986).
22. Fry, B. D., Jannasch, H. W. & Molyneux, L. M. *Deep-Sea Res.* **A38**, S1003–S1010 (1992).
23. Goyet, C. & Bradshaw, P. W. *Deep-Sea Res.* **A38**, S1049–S1058 (1992).
24. Buesseler, K. O., Livingston, H. D., Ivanov, L. & Romanov, A. *Deep-Sea Res.* (in the press).
25. Calvert, S. E. in *Facets of Modern Biogeochemistry* (eds Ittekkot, V. *et al.*) 326–352 (Springer, Berlin, 1990).
26. Tyson, R. V., Wilson, R. C. L. & Downie, C. *Nature* **277**, 377–380 (1979).
27. Dickman, M. & Artuz, I. *Nature* **275**, 191–195 (1978).
28. Hay, B. J. *et al. Deep-Sea Res.* **6**, 911–928 (1990).
29. Hayes, J. M., Freeman, K. H., Popp, B. N. & Homan, C. H. *Org. Geochem.* **16**, 1115–1128 (1990).
30. Jones, G. A. *Eos* **71**, 152 (1990).
31. Calvert, S. E. *et al. Nature* **350**, 692–695 (1991).

ACKNOWLEDGEMENTS. We thank J. J. Middelburg for constructive criticism and the crew of RV *Knorr* for assistance during sampling. B. Hay provided the unit 1 transition sapropel and unit 2 sediments and helped in assigning sediment ages. This work was partly supported by the Netherlands Foundation for Earth Science Research (AWON), by a Pioneer grant to J.S.D. from the Netherlands Organization for Scientific Research (NWO), and by the U.S. NSF and Office of Naval Research.

Entrainment in turbulent gravity currents

Mark A. Hallworth*, Jeremy C. Phillips†, Herbert E. Huppert* & R. Stephen J. Sparks†

* Institute of Theoretical Geophysics, Departments of Earth Sciences and Applied Mathematics and Theoretical Physics, University of Cambridge, 20 Silver Street, Cambridge CB3 9EW, UK

† Department of Geology, University of Bristol, Bristol BS8 1RJ, UK

LABORATORY gravity currents are frequently used to model a range of environmental and industrial flows¹. The manner in which these flows become diluted with distance by the surrounding fluid has important implications for turbidity currents², pyroclastic flows^{3,4}, avalanches⁵, accidental dense gas releases⁶, fire propagation⁷ and emission of industrial pollutants. Here we present an experimental technique for quantifying the entrainment of ambient fluid into the head of a gravity current propagating along a horizontal surface. The technique relies on the neutralization of an alkaline current by entrainment of acidic ambient fluid, and is visualized by using a pH indicator. Dimensional analysis indicates that the proportion of ambient fluid entrained into a gravity current head depends only on the initial volume of the current and distance from the release point, and is independent of the initial value of the density difference. This result is confirmed by the experimental data, which also show that little dilution of the head takes place during the slumping phase^{8,9}. Thereafter the dilution increases with the downstream distance, in quantitative agreement with the predictions of a theoretical model which evaluates the volume of entrained fluid. We apply the results to show that sediment slumps of initially high sediment concentrations will become dilute tur-

bidity currents owing to entrainment of sea water before they have propagated extensively over the floors of ocean basins.

Photographs of a two-dimensional gravity current formed by the sudden release of a fixed volume of dense fluid behind a lock gate into less dense ambient fluid are shown in Fig. 1. Three distinct flow regimes are seen¹: a slumping phase^{8,9}, during which the gravity current head moves at a roughly constant velocity (Fig. 1a); an inertia-buoyancy regime, during which the length of the current increases like the two-thirds power of time¹⁰ (Fig. 1b); and a final regime during which viscous forces dominate over inertial forces¹¹. The point of transition from the slumping phase to the inertia-buoyancy regime has been experimentally related to the fractional depth of the denser fluid initially behind the lock gate⁹. Beyond this point a gravity current has a characteristic form with a roughly hemispherical, lobate head, which entrains ambient fluid by shear instabilities at its trailing edge and by overriding and engulfing ambient fluid at the nose. The volume of the head decreases with distance as fluid is left behind in a thin tail beneath a diffuse mixed zone.

We have determined the proportion of ambient fluid entrained as a function of distance by using a simple neutralization technique. A known amount of sodium hydroxide and some universal pH indicator were mixed into saline solution behind a lock gate placed near one end of a horizontal rectangular channel. The final alkaline solution was purple (pH > 10). The water ahead of the gate was acidified with hydrochloric acid to achieve a value of pH ≤ 4. On release, the dense, alkaline gravity current advanced over the base of the channel, entraining acidic ambient fluid. Turbulence within the head rendered the mixing virtually instantaneous. At a distance from the source which depended on the initial excess concentration of alkali, enough ambient fluid had been entrained to neutralize the gravity current head. At this point an abrupt, homogeneous change in the colour of

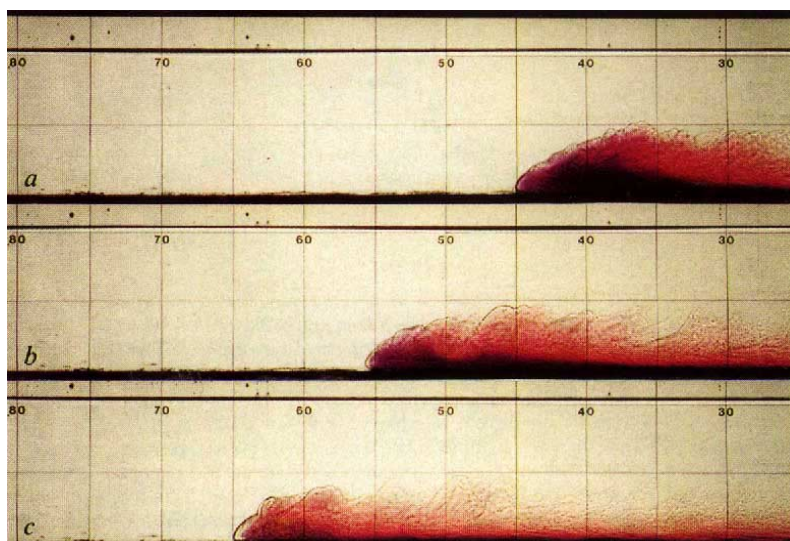


FIG. 1 Photographs of a two-dimensional gravity current formed from the sudden release of a 65 cm³ saline solution with $x_0 = 5.0$ cm, $h_0 = H = 11.3$ cm and $g' = 20$ cm s⁻². a, Slumping phase; b, inertia-buoyancy regime; c, current just at the position at which the head changes colour.

the pH indicator from purple to red was observed throughout the whole head. The fluid composition within the gravity current head at neutrality was determined by prior titration of the initial ambient and dense fluid. Experiments using different initial concentrations of alkali, but otherwise identical initial conditions, mapped out the volumetric proportion of ambient fluid entrained into the gravity current head as a function of distance from the source.

A typical experimental run is shown in Fig. 1. In Fig. 1c the gravity current head has just attained neutrality. The whole gravity current head changes colour rapidly, over less than 2 cm, indicating that it is well mixed. With distance from the lock, the head decreases in volume as fluid is laid down to form the tail of the gravity current. The subsequent velocities in the tail are very much less than that of the head. This contrasts with a current maintained by a continuous flux, for which fluid from the body feeds into the head.

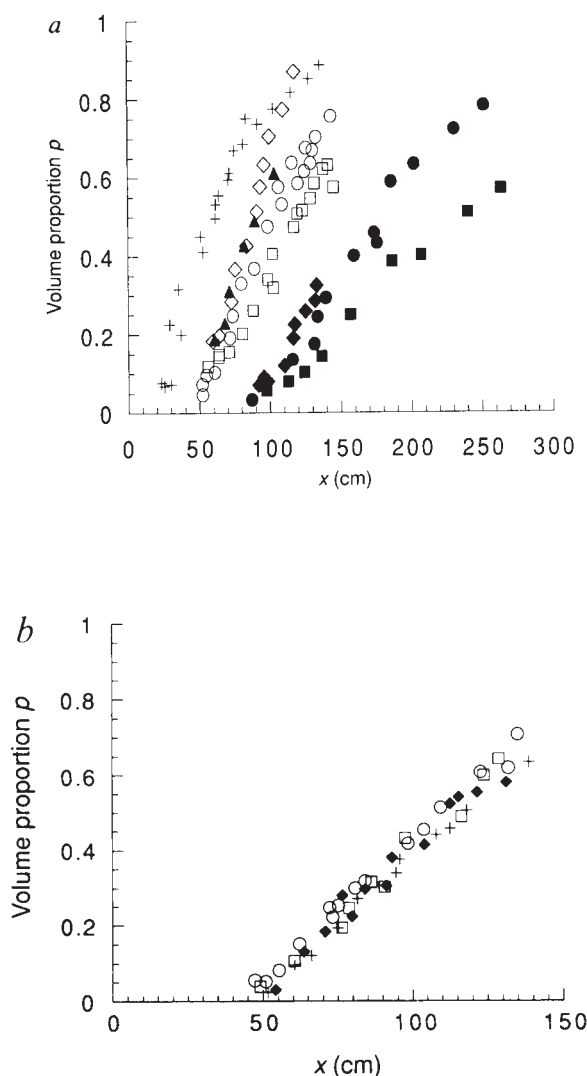


FIG. 2 Measured volume proportion of ambient fluid entrained into the head of a gravity current, p , as a function of the distance from the lock. Each experimental point corresponds to a different run, with conditions as follows: a, $+$ $x_0 = 2.5$ cm, $h_0 = H = 9.3$ cm; $\diamond$ $x_0 = 5.0$ cm, $h_0 = H = 6.0$ cm; $\circ$ $x_0 = 5.0$ cm, $h_0 = H = 9.3$ cm; $\square$ $x_0 = 5.0$ cm, $h_0 = H = 13.0$; $\blacklozenge$ $x_0 = 10$ cm, $h_0 = H = 6.5$ cm; $\bullet$ $x_0 = 10$ cm, $h_0 = H = 13.0$; $\blacktriangle$ $x_0 = 10$ cm, $h_0 = H = 23.0$; $\blacklozenge$ $x_0 = 10$ cm, $h_0 = 5.0$ cm, $H = 35.0$ cm. For all experiments $g' = 20$ cm s⁻² except for the last ($\blacktriangle$), for which $g' = 37$ cm s⁻². b, $x_0 = 5.3$ cm, $h_0 = H_0 = 10$ cm for different values of g' : $\circ$ $g' = 10$ cm s⁻²; $+$ $g' = 20$ cm s⁻²; $\blacklozenge$ $g' = 40$ cm s⁻²; and $\square$ $g' = 80$ cm s⁻².

To quantify the experimental results, we denote by p the fraction of ambient fluid entrained into the gravity current head as defined by

$$p = V_a / (V_a + V_b) \quad (1)$$

where V_a and V_b are the volumes of ambient and original fluid per unit width in the head respectively. The value of p will initially be 0 (no entrainment) and will approach 1 at large distances from the lock, reflecting the high proportion of entrained fluid. Figure 2 presents our measurements of p as a function of x , the distance downstream from the position of the gate, for a range of different initial conditions including three different lock lengths, x_0 . For some distance there is little entrainment into the head. Beyond this distance, p increases significantly downstream. We can identify the beginning of the significant mixing phase with the slumping distance, x_s , empirically determined⁹ as

$$x_s / x_0 = 3 + 7.4 h_0 / H \quad (2)$$

where h_0 is the height of the dense fluid initially behind the lock gate and H is the total depth of the ambient fluid. For experiments with the same lock length, x_0 , the value of p at a fixed value of x decreases monotonically with increasing volume of initial fluid behind the gate.

A systematic collapse of all these data can be obtained by dimensional analysis. We denote by A_0 the initial volume per unit width of the heavy fluid of density $\rho_0 + \Delta\rho$, where ρ_0 is the density of the ambient. Dimensional arguments then suggest that the proportion of ambient fluid entrained into the gravity current head will depend on: A_0 , of dimensions L^2 ; the initial reduced gravity $g'_0 = g\Delta\rho/\rho_0$, of dimensions LT^{-2} ; and the length of the gravity current x , of dimensions L . Because p is dimensionless, it must be independent of g'_0 , the only external parameter dependent on T . This conclusion is verified by the results of experiments at different values of g'_0 (Fig. 2b).

To collapse the data further, we introduce the entrainment ratio or dilution, r , defined by

$$r = p / (1 - p) = V_a / V_b \quad (3)$$

which is the volume ratio of ambient to original fluid in the head, where $r = 0$ initially ($x = 0$) and tends to infinity for long currents ($x \rightarrow \infty$). As both p and $r = 0$ for $0 < x < x_s$, we introduce the new variable $y = x - x_s$, whereupon dimensional analysis suggests that r is a function of y^2/A_0 only. Replotting all our data in this way, we obtain the satisfactory collapse shown in Fig. 3.

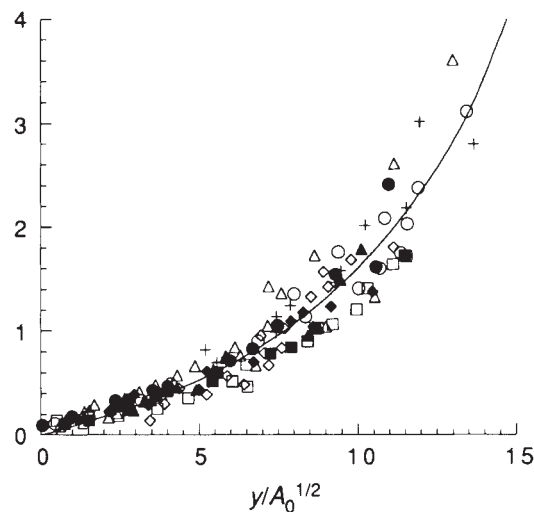


FIG. 3 Entrainment ratio r as a function of $y/A_0^{1/2}$ for all our data, with the best-fit theoretical curve $r = (1 - 0.0347 y/A_0^{1/2})^{-2.236} - 1$.

The form taken by the data can be explained as follows. The volume of the head per unit width beyond the slumping phase, V , increases by entrainment of ambient fluid and simultaneously decreases by leaving fluid behind in the tail. These processes can be represented, respectively, by

$$\frac{dV}{dy} = \alpha V^{1/2} - k V^{1/2} \quad (4a)$$

$$V(0) = A_0 \quad (4b)$$

where α is an entrainment constant and the constant k is representative of the ratio of the height of the tail to that of the head. The expression $V^{1/2}$ indicates that the spatial rate of change of volume is proportional to the height of the head; and this is the only form possible from dimensional considerations. The solution to equation (4) is

$$V/A_0 = [1 - \frac{1}{2}(k - \alpha)y/A_0^{1/2}]^2 \quad (5)$$

This expression indicates that the head considered as a separate entity runs out of volume at $y = y_c \equiv 2A_0^{1/2}/(k - \alpha)$. Thereafter the fluid must continue to move forward, driven by a horizontal density gradient within the tail. In most (but not all) natural and laboratory situations, viscous effects dominate before the current has reached $y = y_c$.

Keeping account of ambient and original fluid individually, we find that the volume of ambient and original fluid in the head are described by

$$\frac{dV_a}{dy} = \alpha V^{1/2} - kpV^{1/2}, \quad \frac{dV_b}{dy} = -k(1-p)V^{1/2} \quad (6a, b)$$

$$V_a(0) = 0, \quad V_b(0) = A_0 \quad (6c, d)$$

Substituting $p = V_a/V$ and $1-p = V_b/V$, with V represented by equation (5), into equation (6) and integrating the result, we obtain

$$r = [1 - \frac{1}{2}(k - \alpha)y/A_0^{1/2}]^{-2\alpha/(k - \alpha)} - 1 \quad (7)$$

which becomes infinite as y approaches y_c . The values of α and k that give a best fit to all our data are $\alpha = 0.078$ and $k = 0.147$ ($k - \alpha = 0.069$), with a computed error of $\sim 3\%$. Equation (7), incorporating these values, is displayed in Fig. 3 along with all the data. The agreement between the form of the curve and the data is good. The total volume per unit width of entrained plus original fluid left behind in the tail, V_T , is given by

$$V_T = k \int_0^{y_c} V^{1/2} dy \quad (8a)$$

$$= k/(k - \alpha)A_0 = 2.12A_0 \quad (8b, c)$$

so the total volume of fluid entrained into the gravity current head is slightly in excess of the original volume. The average dilution is therefore just above one, although the local dilution can be much larger.

The results apply quantitatively only to fixed volume releases of fluid in a constrained channel. However, entrainment rates are likely to be similar in more complex flow geometries or flow conditions. Our results demonstrate that entrainment is substantial and that any natural current will become highly diluted downstream. One application is to turbidity currents where the density difference arises from suspended particles. In ocean basins, large turbidity currents are typically generated by slumps of sediment on the upper continental slope¹²⁻¹⁴. Individual slumps may have volumes ranging from less than 0.1 km^3 to several tens of cubic kilometres. They must initiate as flows with high concentrations of suspended sediments. On the assumption that most currents are erosive in the feeder channels or canyons, entrainment into the flow head once deposition has begun will rapidly cause the head of the current to evolve toward low concentration. For example, with values of a length of 10 km

along the axis of the feeder canyon and a height of 100 m, $A_0 = 10^6 \text{ m}^2$. Inserting this value into equation (7), and extrapolating our concepts to more complex natural situations, we find that the head of a current initiated in water 2 km deep with an initial concentration of 40% would be diluted to $\sim 8\%$ over a distance of 15 km (to 3% over 20 km) beyond a slumping distance of ~ 33 km, and would cease propagating in the form of a classical gravity current in ~ 62 km from the source. We would expect the resulting deposits to evolve from those characteristic of high concentration close to the canyon mouth to those characteristic of low concentration further away. This prediction agrees with general observations². Turbidity currents are much more complex than simple laboratory currents because of effects that include lateral spreading, variations in fluid and particle fluxes, sedimentation and erosion. Nevertheless this calculation shows that entrainment is an effective way of evolving an initial high-concentration sediment slump into a dilute turbidity current. $\square$

Received 19 January; accepted 17 March 1993.

1. Simpson, J. E. *Gravity Currents in the Environment and the Laboratory* (Halsted, Chichester, 1987).
2. Allen, J. R. L. *Principles of Physical Sedimentology* (Allen & Unwin, London, 1985).
3. Wilson, C. J. N. *Sci. prog. Oxf.* **70**, 171-207 (1986).
4. Sparks, R. S. J. et al. *Earth planet. Sci. Lett.* **114**, 243-257 (1993).
5. Hopfinger, E. J. *A. Rev. Fluid Mech.* **15**, 47-76 (1983).
6. Britter, R. E. *A. Rev. Fluid Mech.* **21**, 317-344 (1989).
7. Emmons, H. W. *Proc. 3rd Int. Symp. Fire Safety Science* (eds Cox, G. & Langford, B.) 249-260 (Elsevier, London, 1991).
8. Huppert, H. E. & Simpson, J. E. *J. Fluid Mech.* **99**, 785-799 (1980).
9. Rottman, J. W. & Simpson, J. E. *J. Fluid Mech.* **135**, 95-110 (1983).
10. Hout, D. P. *A. Rev. Fluid Mech.* **4**, 341-368 (1972).
11. Huppert, H. E. *J. Fluid Mech.* **121**, 43-58 (1982).
12. Piper, D. J. W., Shor, A. N., Farre, J. A., O'Connell, S. & Jacobi, R. *Geology* **13**, 538-541 (1985).
13. Heezen, B. C. & Ewing, M. *Am. J. Sci.* **250**, 849-873 (1952).
14. Heezen, B. C., Erickson, D. B. & Ewing, M. *Deep-Sea Res.* **1**, 193-202 (1954).

ACKNOWLEDGEMENTS. We thank R. E. Britter and E. J. Hinch for discussions and R. Bonnetaze, B. Dade, R. Kerr, P. Linden, J. Lister, N. McCave, H. Pantin and J. S. Turner for comments on an earlier version. This research is partially supported by Venture Research International.

Relationship between eruption volume and neodymium isotopic composition at Unzen volcano

Chang-Hwa Chen^{*†}, Donald J. DePaolo^{*},
Setsuya Nakada[‡] & Yuch-Ning Shieh[§]

^{*} Center for Isotope Geochemistry, Department of Geology and Geophysics, University of California and Earth Sciences Division, Lawrence Berkeley Laboratory, Berkeley, California 94720, USA
[†] Institute of Earth Sciences, Academia Sinica, PO Box 23-59, Taipei, Taiwan 10764

[‡] Department of Earth and Planetary Sciences, Kyushu University, Fukuoka 812, Japan

[§] Department of Earth and Atmospheric Sciences, Purdue University, West Lafayette, Indiana 47907

SILICA-rich lavas, erupted at island-arc or continental volcanoes, are often produced by a complex process involving the assimilation of crust into a crystallizing, mantle-derived basaltic magma¹. The different strontium, neodymium and oxygen isotopic compositions of mantle-derived magmas and continental crust provide a powerful method for tracing the different contributions to continental silicic magmas, and for understanding the parameters controlling the composition and volume of erupted magma¹⁻⁴. In the large rhyolite eruptive centres of the western United States, the largest-volume, explosive rhyolite eruptions have more mantle-like Nd isotope ratios than other silicic lavas from the same centre²⁻⁴, a relationship that has been interpreted as reflecting increased influx of

TABLE 1 Isotopic and chemical data for lava samples of Unzen volcano and surrounding area

Unit Age Lithology	Fugendake				Kusenbudaka		Takadake	
	AD 1792 Dacite	AD 1663 Andesite	4 kyr Dacite	5 kyr Dacite	100–175 kyr Andesite	100–175 kyr Dacite	180–275 kyr Dacite	180–275 kyr Dacite
SiO ₂	66.27	57.64	62.23	65.16	58.13	65.39	64.33	64.48
TiO ₂	0.65	1.01	0.74	0.68	0.87	0.62	0.70	0.79
Al ₂ O ₃	15.79	17.51	16.39	16.14	16.97	15.87	16.35	16.39
Fe ₂ O ₃	4.29	7.50	5.47	4.84	7.26	4.79	5.04	5.41
MnO	0.09	0.13	0.10	0.09	0.12	0.09	0.10	0.09
MgO	2.24	4.26	2.72	2.31	3.92	2.39	2.17	2.81
CaO	4.57	7.29	5.44	4.57	6.73	4.14	4.61	3.93
Na ₂ O	3.79	3.20	3.55	3.66	3.32	3.56	3.65	3.46
K ₂ O	2.40	1.72	2.26	2.67	1.80	2.57	2.43	2.73
Sr	278	364	322	314	471	361	335	335
Rb	78	54	70	85	46	78	77	52
δ ¹⁸ O (‰)	8.9	8.1	9.1	9.0	8.9	9.4	9.4	9.1
⁸⁷ Sr/ ⁸⁶ Sr	0.70462	0.70416	0.70454	0.70442	0.70419	0.70479	0.70479	0.70464
ε _{Nd}	0.58	2.43	1.93	1.52	2.07	0.58	0.46	1.49

Unit Age Lithology	Pre-Unzen				Present eruption					Northwest
	500 kyr Andesite	1.42 Myr Andesite	600 kyr Basalt	1.11 Myr Basalt	AD 1991 May Basalt	AD 1991 September Dacite	AD 1991 Xenolith	AD 1992 March Dacite	AD 1992 April Dacite	Kyushu Xenolith
SiO ₂	56.60	57.24	50.29	50.68	64.52	65.08	52.50	64.82	65.02	48.96
TiO ₂	1.01	1.01	1.39	1.61	0.68	0.64	1.20	0.66	0.66	0.15
Al ₂ O ₃	17.39	17.38	16.05	15.86	16.01	15.94	18.00	15.85	15.84	16.35
Fe ₂ O ₃	8.09	7.77	10.38	9.64	4.84	4.66	10.20	4.69	4.65	5.42
MnO	0.11	0.17	0.18	0.18	0.06	0.06	0.20	0.08	0.08	0.10
MgO	4.35	3.26	7.62	7.02	2.54	2.42	4.46	2.37	2.41	13.33
CaO	7.34	6.89	9.47	8.87	4.97	4.81	8.26	4.93	4.79	13.88
Na ₂ O	3.14	3.53	2.87	3.14	3.79	3.77	3.36	3.95	3.83	1.34
K ₂ O	1.62	1.75	1.06	1.43	2.42	2.45	1.35	2.49	2.56	0.07
Sr	509	486	420	452	320	318				
Rb	46	55	23	28	84	85				
δ ¹⁸ O (‰)	9.1	8.6	7.5	7.6	8.3	8.0	7.4			6.9
⁸⁷ Sr/ ⁸⁶ Sr	0.70453	0.70409	0.70395	0.70437	0.70449	0.70446	0.70435	0.70438	0.70448	0.70416
ε _{Nd}	1.41	2.67	2.92	3.45	2.00	1.59	2.73			4.34

* Major element oxide concentrations are given as weight per cent. Sr and Rb concentrations are in units of µg per g (p.p.m.). Oxygen isotope values (‰) are given as $\delta^{18}\text{O} = (^{18}\text{O}/^{16}\text{O})_{\text{sample}} / (^{18}\text{O}/^{16}\text{O})_{\text{SMOW}} - 1$ where SMOW is standard mean ocean water; analytical uncertainty is ± 0.2 . The $^{87}\text{Sr}/^{86}\text{Sr}$ values are normalized to $^{86}\text{Sr}/^{88}\text{Sr} = 0.1194$ to correct for mass discrimination; analytical uncertainty is ± 0.00002 . The ϵ_{Nd} values represent $^{143}\text{Nd}/^{144}\text{Nd}$ ratios normalized to a standard chondritic value: $\epsilon_{\text{Nd}} = [^{143}\text{Nd}/^{144}\text{Nd}_{\text{measured}} / ^{143}\text{Nd}/^{144}\text{Nd}_{\text{chondritic}} - 1] \times 10^4$. $^{143}\text{Nd}/^{144}\text{Nd}_{\text{chondritic}}$ is taken to be 0.511836, and the correction for mass discrimination is done assuming $^{146}\text{Nd}/^{142}\text{Nd} = 0.636151$. Analytical uncertainty in ϵ_{Nd} is ± 0.3 .

mantle-derived basaltic magma to a crustal magma chamber before large-volume eruptions¹. Here we report isotope data for lavas from Unzen volcano, which suggest a similar relationship: the Nd isotope composition is more mantle-like in three larger-volume dacite eruptions (>0.1 km³) than in one small-volume (0.02 km³) eruption. We accordingly suggest that, in small-volume systems like Unzen, where the timescales for magma-chamber evolution are of the order of decades, isotope data such as those presented here might be used in volcanic hazard evaluation.

Unzen volcano⁵ began erupting explosively in 1991 and has taken 43 lives. The eruptive history of the volcano extends over about 275,000 years and is divided into three main stages^{6,7}: the Fugendake (0–125 kyr ago), Kusenbudake (100–175 kyr), and Takadake (175–275 kyr). The three eruptive stages form three volcanic bodies in the central Simabara peninsula, each composed of hornblende andesite and dacite lava flows and pyroclastic deposits. The most recent eruptions preceding the present eruption occurred in AD 1663 (andesite) and AD 1792 (dacite). The present eruption began as phreatic activity in November 1990; explosive activity commenced in May 1991 and continues⁸.

We collected 18 samples from the three main eruptive episodes of the Unzen volcano as well as pre-Unzen lavas from the vicinity. One basaltic xenolith sample from northwestern Kyushu is also included. The ages of the samples span from 1.42 Myr ago to the 1991–92 eruption^{6,7}. The samples (Table 1) were analysed for major elements and for the isotopic compositions of strontium, neodymium and oxygen using established techniques^{9,10}.

There is some correlation between ϵ_{Nd} and the silica content of the lavas (Fig. 1), and there is good correlation between the ϵ_{Nd} values and the $\delta^{18}\text{O}$ values (see Table 1 for definitions).

The figure shows schematically the expected composition of typical island-arc basaltic rocks of the western Pacific (denoted 'M' for mantle composition^{11,12}; and the typical rocks of the continental-type crusts of southeastern Japan ('Crust')^{13,14}. All Unzen lavas have isotopic compositions intermediate between mantle- and crustal-type values. Basaltic lavas have isotopic parameters closest to the mantle values; lavas with higher silica contents have values closer to crustal values. The data indicate that lavas of Unzen volcano are mixtures of magma from both the mantle and the crust. None of the lavas is purely derived from the mantle, and none of the lavas, including the dacites, is purely derived from the crust. The pattern of Fig. 1a is a strong indication that the higher-silica lavas are formed by fractional crystallization from lower-silica lavas¹⁵, and that crustal material with low- ϵ_{Nd} and high $\delta^{18}\text{O}$ is added as fractional crystallization proceeds. This process is known as AFC (assimilation-fractional crystallization).

The isotopic compositions of the dacite lavas are separable into two groups primarily on the basis of the ϵ_{Nd} values (Fig. 2). Volume data are available for four of the most recent eruptions of dacite (Fig. 3). The volume of the older eruptions cannot be estimated with confidence because the lavas are both partially covered by younger units and partially eroded. The group-1 lavas, characterized by ϵ_{Nd} values of +1.5 to +2.0, include three eruptions of relatively large volume (0.1 to 0.5 km³). The group-2 lavas, which have ϵ_{Nd} values of +0.6, include the one documented small-volume dacite eruption (1792; 0.02 km³). The volume of the present eruption is currently 0.11 km³. The limited data suggest a relationship between ϵ_{Nd} value and eruption volume when comparing differentiated lavas of similar major element composition.

The Unzen data are too few to demonstrate a consistent relationship between ϵ_{Nd} and eruption volume. However, we consider the observation noteworthy because an analogous pattern is observed for Pleistocene rhyolite caldera systems of the western United States. In the Jemez system of New Mexico, andesite and rhyolite domes, flows and tuffs erupted over the period from 8 Myr to <1 Myr ago have ϵ_{Nd} values of -2.5 to -5.5 (ref. 4). The much larger rhyolite ashflow eruption at 1.4 Myr ago has an ϵ_{Nd} value of 0, and the rhyolite domes that preceded it by a few hundred thousand years have values of -0.5 to -1.5 (ref. 4). Similarly, the ϵ_{Nd} values of rhyolite flows and tuffs at Long Valley, California, in the time period 2.0–1.2 Myr ago are -3 to -4, whereas the ϵ_{Nd} values of the much larger Bishop Tuff eruption (0.76 Myr) and the rhyolites that

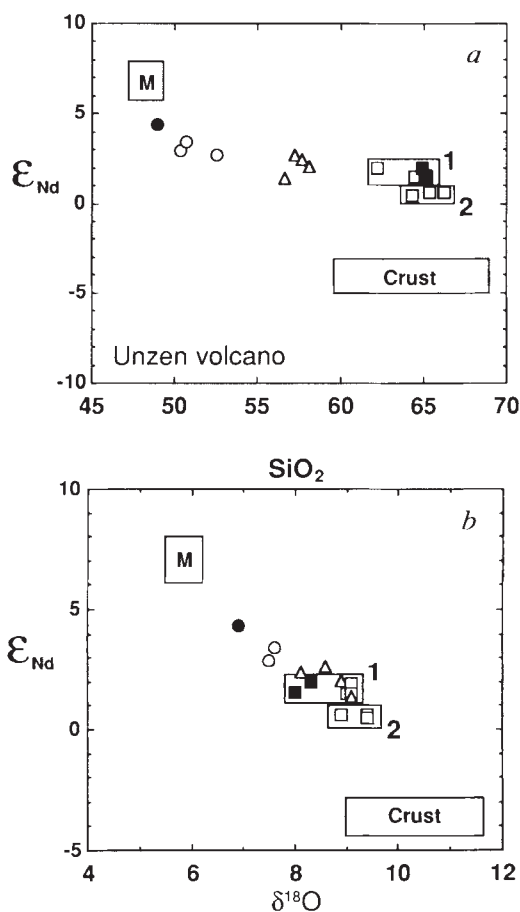


FIG. 1 *a*, Plot of ϵ_{Nd} against silica content for all samples measured from Unzen volcano and surrounding area. Fields labelled 'M' and 'Crust' are the inferred compositions of parental mantle-derived basalt and the typical contaminant to this basalt from the crust. Square symbols represent dacite samples, triangles are andesite; large circles are basalt and basaltic andesite. Filled square symbols are 1991 dacite. The position and shape of the data array indicate that the more silica-rich lavas are derived from basalt by a process that involves admixing crustal rock material (which lowers ϵ_{Nd} and increases SiO_2 simultaneously), and fractional crystallization (which increases SiO_2 but does not affect ϵ_{Nd}). The filled circle represents a basaltic xenolith from northwest Kyushu (Table 1) which has isotopic characteristics similar to those inferred for the parental magma of the Unzen volcano. Dacite samples can be separated into two groups on the basis of ϵ_{Nd} . *b*, Plot of ϵ_{Nd} against $\delta^{18}\text{O}$ for the Unzen volcano samples. Symbols are the same as in *a*. This data array also indicates that crustal rock material is mixed with mantle magma as fractional crystallization proceeds, and that even the low- SiO_2 magmas are contaminated with a substantial amount of crustal rock material. Group 1 and 2 dacites are indicated.

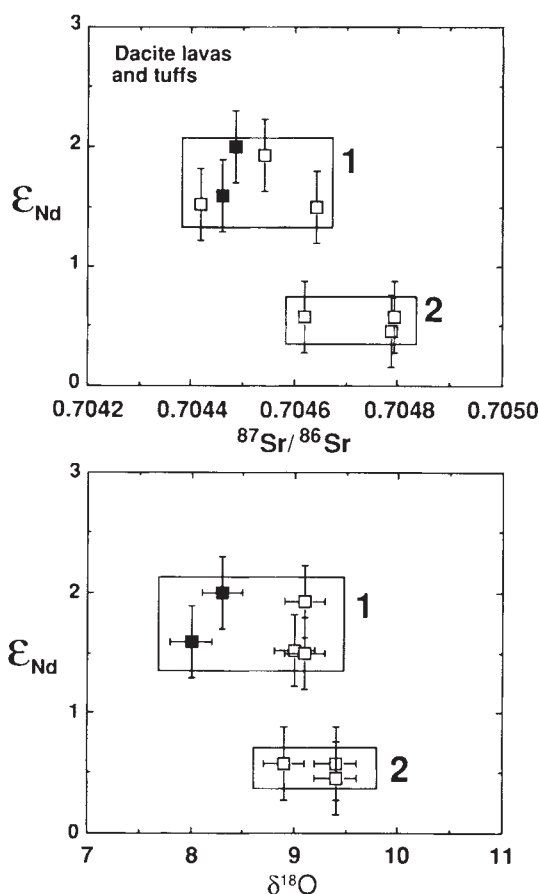


FIG. 2 Expanded view of ϵ_{Nd} , $^{87}\text{Sr}/^{86}\text{Sr}$ and $\delta^{18}\text{O}$ variations in dacite of Unzen volcano. In both cases, and in the ϵ_{Nd} - SiO_2 plots as well (Fig. 1*a*), there is correlation between the two parameters. Only ϵ_{Nd} appears to be consistently different between two populations (1 and 2) of lavas; for the other parameters there is overlap.

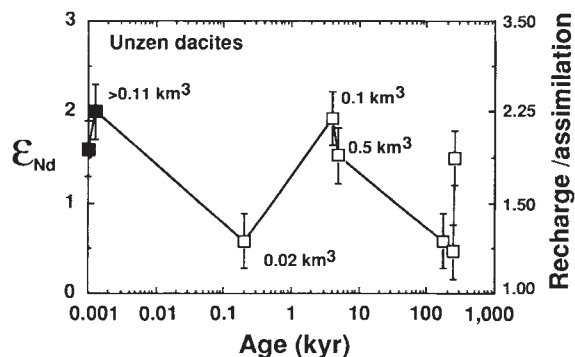


FIG. 3 ϵ_{Nd} against age for dacite lavas of Unzen volcano. The volumes for the eruptions are labelled where known. Volume is difficult to determine for the older lavas because they are largely covered and/or eroded. Three of the four high- ϵ_{Nd} eruptions are known to have large volumes. The 1792 eruption had a much smaller volume and a lower ϵ_{Nd} . Calculated steady-state values of r/a (recharge/assimilation ratio) are shown on the scale at right; the values are calculated using $\epsilon_r = +6$ (value for the recharging basalt), $\epsilon_a = -4$ (value for the crust) and $C_a/C_r = 1.5$ (ratio of Nd concentration in crustal assimilate to Nd concentration in recharging basalt). Steady state $r/a = (C_a/C_r) (\epsilon_a - \epsilon_m) / (\epsilon_m - \epsilon_r)^4$, where ϵ_m is the value for the resulting magma (the measured values for the dacites in this case).

preceded it by a few hundred thousand years are near -1 (ref. 2). Similar features are also evident in time-series data from the Yellowstone volcanic system, although they have been interpreted differently³.

A common feature of volcanic systems exhibiting a relationship between eruption volume and ϵ_{Nd} value is that assimilation of crustal material by mantle magma is evident¹⁻⁴. Our preferred model to explain the relationship^{1,4,15} requires the magmatic evolution process to include recharge of the system with mantle-derived basalt as well as fractional crystallization and assimilation, and is referred to as RAFC¹⁶. A primary requirement of any model is that it account for the observation that ϵ_{Nd} fluctuates between high and low values in each volcanic centre (as in Fig. 3).

In the RAFC model, the ϵ_{Nd} value of a magma chamber is primarily determined by the ratio of recharge to assimilation (r/a)¹⁶. The inflation rate of the magma system (dM/dt), however, depends on the recharge rate (m_r) plus the assimilation rate (m_a), minus the rate of crystallization (m_c)

$$dM/dt = m_a + m_r - m_c$$

where the italicized parameters are rates in mass per unit time. The model holds that m_a and m_c tend to be fixed (or nearly so) by the ambient conditions^{1,15}, mainly by the temperature of the rock surrounding the magma system in the crust. The recharge rate, however, is controlled by processes in the mantle and can vary with time over a wide range. The inflation rate of the crustal magma system is fastest when m_r is largest, which is also when r/a is largest. When the magma body reaches its largest size, potentially leading to a large eruption, the ϵ_{Nd} value in the magma chamber tends to be higher (closer to the mantle-type value) than at other times. The contemporaneity of high r/a and high ϵ_{Nd} depends on the residence time of Nd in the chamber; however, during periods of inflation the effective residence time is short¹⁷ and ϵ_{Nd} tracks r/a closely. The size of the magma body is given by the time-integral of the inflation rate. Accumulation of large volumes of differentiated magma like dacite requires an elevated recharge rate maintained over a sufficient time period. For high recharge rates over short times, there is little assimilation or crystallization, and mafic magma is erupted instead of silicic magma; this yields the observed correlation between magma composition (for example the silica content) and ϵ_{Nd} (Fig. 1a). This model, which is based on isotopic data, differs markedly from others¹⁸ which require *a priori* that silicic magmas be melted entirely from the crust.

Our results suggest that ϵ_{Nd} of early or precursory eruptive products could be a qualitative indicator of the maximum size of a continuing or impending eruption. The relationship between ϵ_{Nd} and eruption volume is not linear; a shift of 1.0 to 1.5 units of ϵ_{Nd} corresponds to about one order of magnitude in eruption volume, and is not strictly monotonic because factors other than recharge rate may also affect eruption volume. The Unzen samples we measured were collected in 1991 at a time when the total erupted volume was $\sim 0.03 \text{ km}^3$, but correspond in ϵ_{Nd} to eruptions of $> 0.1 \text{ km}^3$ volume. In comparison with large-volume rhyolitic centres, the timescales for magma chamber inflation, precursory activity and eruption at Unzen are orders of magnitude smaller, and closer to timescales of human concern¹⁹.

The sensitivity of Nd isotopes for assessing potential eruption size depends on two factors: the minimum value of r/a that is possible for the system, and the isotopic contrast between the mantle and crustal components of the system^{1,15}. In the case of Unzen, the isotopic contrast between crust and mantle is moderate, but the amount of crustal assimilation is relatively great (Fig. 1), consistent with the observed high geothermal gradient^{20,21}. In island arcs where there is no underlying continental-type crust, there may be no Nd isotopic contrast between crust and mantle and therefore no possibility of using Nd isotopic measurements in the manner described here. In

primarily basaltic volcanoes, assimilation may not play as large a role, and different behaviour can be expected^{22,23} □

Received 13 October 1992; accepted 8 March 1993.

- DePaolo, D. J., Perry, F. & Baldrige, W. S. *Proc. R. Soc. Edinb.* **83**, 439-446 (1992).
- Hildrey, A. N. *et al. Earth planet. Sci. Lett.* **94**, 274-290 (1989).
- Hildrey, W., Halliday, A. N. & Christiansen, R. L. *J. Petrol.* **32**, 63-138 (1991).
- DePaolo, D. J., Perry, F. V. & Baldrige, W. S. *Geol. Soc. Am. Abstr.* **23**, A396 (1991).
- Letouzey, J. & Kimura, M. *Mar. Petrol. Geol.* **2**, 111-130 (1985).
- Nakada, S. & Kamata, H. *Bull. Volcanol. Soc. Japan* **33**, 273-289 (1988).
- Nakada, S. & Tanaka, M. *Bull. Volcanol. Soc. Japan* **36**, 113-121 (1991).
- Nakada, S. & Fujii, T. *J. Volcanol. geotherm. Res.* (in the press).
- DePaolo, D. J. *J. geophys. Res.* **86**, 10470-10488 (1981).
- Clayton, R. N. & Mayeda, T. K. *Geochim. cosmochim. Acta* **27**, 43-52 (1963).
- DePaolo, D. J. *Neodymium Isotope Geochemistry: An Introduction* (Springer, Heidelberg, 1988).
- Chen, C.-H. thesis, National Taiwan Univ. (in Chinese) (1989).
- Matsuhashi, Y. & Kurasawa, H. *J. Volcanol. geotherm. Res.* **18**, 483-510 (1983).
- Terakado, Y. & Nakamura, N. *Contrib. Miner. Petrol.* **87**, 407-417 (1984).
- Perry, F. V., DePaolo, D. J. & Baldrige, W. S. *Geol. Soc. Am. Bull.* (in the press).
- DePaolo, D. J. *J. Petrol.* **26**, 925-951 (1985).
- Albarede, F. *Geochim. cosmochim. Acta* (in the press).
- Huppert, H. E. & Sparks, R. S. J. *J. Petrol.* **29**, 599-624 (1988).
- Spera, F. J. & Crisp, J. A. *J. Volcanol. geotherm. Res.* **11**, 169-187 (1981).
- Okubo, Y., Graf, R. J., Hansen, R. O., Ogawa, K. & Tsui, H. *Japan. Geophys.* **53**, 481-494 (1985).
- Ehara, S. *Tectonophysics* **159**, 269-278 (1989).
- Cortini, M. & Hermes, O. D. *Contrib. Miner. Petrol.* **77**, 47-55 (1981).
- Cortini, M. & Scandone, R. *J. Volcanol. geotherm. Res.* **12**, 393-400 (1982).

ACKNOWLEDGEMENTS. This research was supported by the NSF (D.J.D.) and by the Director, Office of Energy Research, Office of Basic Energy Sciences, Engineering and Geosciences Division of the US Department of Energy.

Ferrous iron oxidation by anoxygenic phototrophic bacteria

Friedrich Widdel*, Sylvia Schnell†, Silke Heising†, Armin Ehrenreich*, Bernhard Assmus* & Bernhard Schink†‡

* Max-Planck-Institut für Marine Mikrobiologie, Fahrenheitstrasse 1, D-2800 Bremen, Germany

† Fakultät für Biologie der Universität Konstanz, Postfach 5560, D-7750 Konstanz, Germany

NATURAL oxidation of ferrous to ferric iron by bacteria such as *Thiobacillus ferrooxidans* or *Gallionella ferruginea*¹, or by chemical oxidation^{2,3} has previously been thought always to involve molecular oxygen as the electron acceptor. Anoxic photochemical reactions⁴⁻⁶ or a photobiological process involving two photosystems⁷⁻⁹ have also been discussed as mechanisms of ferrous iron oxidation. The knowledge of such processes has implications that bear on our understanding of the origin of Precambrian banded iron formations¹⁰⁻¹⁴. The reducing power of ferrous iron increases dramatically at pH values higher than 2-3 owing to the formation of ferric hydroxy and oxyhydroxy compounds^{1,2,15} (Fig. 1). The standard redox potential of $\text{Fe}^{3+}/\text{Fe}^{2+}$ ($E_0 = +0.77 \text{ V}$) is relevant only under acidic conditions. At pH 7.0, the couples $\text{Fe}(\text{OH})_3/\text{Fe}^{2+}$ ($E'_0 = -0.236 \text{ V}$) or $\text{Fe}(\text{OH})_3 + \text{HCO}_3^-/\text{FeCO}_3$ ($E'_0 = +0.200 \text{ V}$) prevail, matching redox potentials measured in natural sediments^{9,16,17}. It should thus be possible for Fe(II) around pH 7.0 to function as an electron donor for anoxygenic photosynthesis. The midpoint potential of the reaction centre in purple bacteria is around $+0.45 \text{ V}$ (ref. 18). Here we describe purple, non-sulphur bacteria that can indeed oxidize colourless Fe(II) to brown Fe(III) and reduce CO_2 to cell material, implying that oxygen-independent biological iron oxidation was possible before the evolution of oxygenic photosynthesis.

Samples of black, reduced sediment stored anoxically in the light had rusty depositions on the glass wall after four weeks without showing development of algae or cyanobacteria. To study this phenomenon in more detail, enrichment cultures were

‡ To whom correspondence should be addressed.

set up under nitrogen gas in butyl rubber-sealed glass bottles with strictly anoxic, bicarbonate-buffered (20 mM) mineral medium¹⁹, pH 7.0, free of sulphide, but containing vitamins and trace elements²⁰. Depending on the inoculum source, either freshwater medium or saltwater medium (with 20.0 g NaCl and 3.0 g MgCl₂·6H₂O per litre) was used. FeSO₄ was added as the sole reductant to the medium to obtain a concentration of 2 or 10 mM, from an anoxic stock solution kept under nitrogen gas. The Fe²⁺ ions reacted, probably with carbonate and phosphate in the medium, to form a fluffy, white precipitate. Cultures were inoculated with few drops of water from the sediment surface of freshwater ponds or from the Jadebusen, North Sea.

After 2–8 weeks of incubation at low light intensity (25 W light bulb, 30–40 cm distance) with or without a long-wavelength (>780 nm) light filter, the precipitate in the bottles turned greenish-grey and then rusty-brown at the sites of highest light intensity. This material consisted of Fe(III) oxides containing embedded cells which did not exhibit fluorescence when checked at 436–520 nm excitation wavelength, indicating the absence of chlorophyll *a*. In a control experiment, the light-dependent oxidation of ferrous iron was not inhibited by addition of 100 µM dichlorophenyldimethylurea which at this concentration is more than sufficient to inhibit photosystem II (ref. 21). After three subcultures, the bacteria were purified in agar dilutions¹⁹.

Three major morphological types were isolated from freshwater sources. One of these, after growth with succinate or acetate, resembled *Rhodomicrobium vannielii* (Fig. 2a) and exhibited a cell cycle typical of this species. After growth with ferrous iron, the cells were covered with highly refractile crusts of Fe(III) precipitates (Fig. 2b). The second type had ovoid to rod-shaped cells similar to *Rhodospseudomonas palustris*. The third cell type contained gas vesicles and morphologically resembled *Thiodictyon* spp., but did not grow with sulphide. These cells were less closely associated with the precipitate (Fig. 2c). Authenticated *Thiodictyon* species²² were also tested; however, they did not oxidize ferrous iron. The strains isolated from a marine enrichment culture grow poorly in the iron medium.

TABLE 1 Quantification of anaerobic phototrophic ferrous iron oxidation by purple bacteria

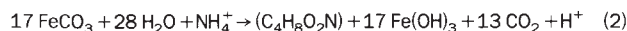
	<i>Rhodomicrobium</i> -like isolate	Gas vesicle-containing isolate
Fe(II) at the beginning (mmol l ⁻¹)*	1.94	10.0
Fe(II) at the end (mmol l ⁻¹)*	0.77	0.5
Fe(II) oxidized (mmol l ⁻¹)	1.17	9.5
Cell mass produced (mg l ⁻¹)	5.2	72
Theoretical cell mass production (mg l ⁻¹)‡		
from equation (1)	8.8	71
from equation (2)	7.0	57

Cultures were incubated for 4 weeks or 12 days, respectively.

* Fe(II) was determined photometrically²⁵ after dissolution of the total precipitate of a culture in 2 M HCl. Fe(III) by the same method after reduction with hydroxylamine²⁶.

† Cell mass was determined by protein. Protein was centrifuged from the HCl-treated culture after addition of 0.6 M (final concentration) trichloroacetic acid, dissolved in 0.1 M NaOH and determined with the biuret reagent²⁷. The cell mass/protein ratio was measured in a separate experiment with cells grown on acetate and then dried to constant weight at 70 °C.

‡ Theoretical cell mass production (approximate formula, CH₂O or C₄H₇O₂N) with ferrous iron as electron donor was calculated by the following equations:



Thus, 1 mmol Fe(II) theoretically yields 7.51 or 6.01 mg cell mass, respectively.

The measured electron balance of iron oxidation versus CO₂-dependent cell mass formation by the freshwater isolate came close to the theoretically expected values (Table 1). No iron oxidation occurred in the dark or in a sterile control experiment: the precipitate remained colourless and the ferrous iron content

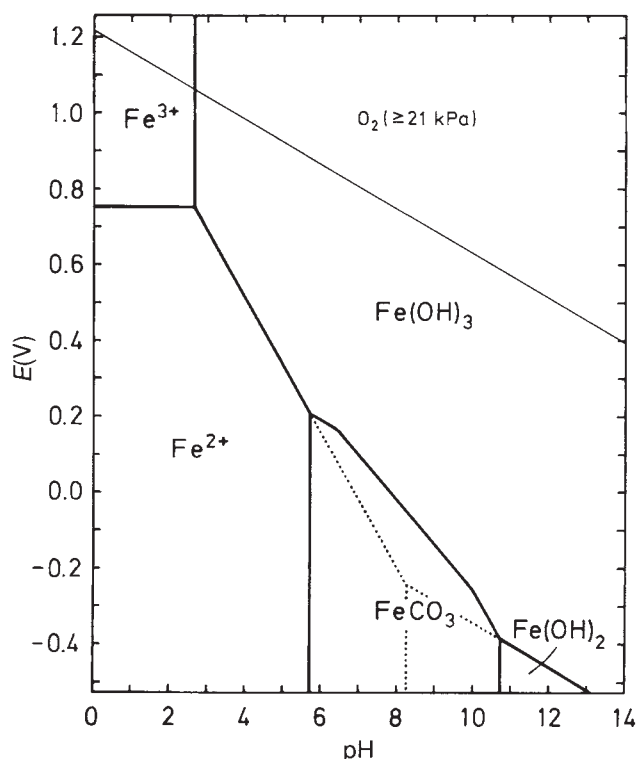


FIG. 1 Graph of *E* versus pH for iron species and O₂ at atmospheric concentration (H₂O below respective line) at 25 °C. The calculations are based on concentrations of 1 × 10⁻³ mol l⁻¹ total iron, 2 × 10⁻² mol l⁻¹ dissolved inorganic carbon (CO₂ + HCO₃⁻ + CO₃²⁻), and non-reacting ions to make up an ionic strength of 0.05; the individual (calculated) activity coefficients of Fe³⁺, Fe²⁺ and HCO₃⁻ are 0.24, 0.48 and 0.81, respectively². Of various solubility products given for Fe(OH)₃, a value of 10⁻³⁸ was taken which refers to a fresh precipitate². On a line demarcating Fe²⁺ or Fe³⁺, the respective ion accounts for 50% (=0.5 × 10⁻³ mol l⁻¹) of the total iron. Dotted lines indicate the extended stability regions of Fe²⁺ and Fe(OH)₂ in the absence of inorganic carbon. For easy overview, H₂ (0 V at pH 0, -0.414 V at pH 7; at 101 kPa), magnetite and Fe(OH)₂²⁺ are not included; hence, the iron species indicated are metastable in certain regions. Metallic iron (below -0.547 V) is outside the panel. Other details and overviews are given in the literature¹⁵.

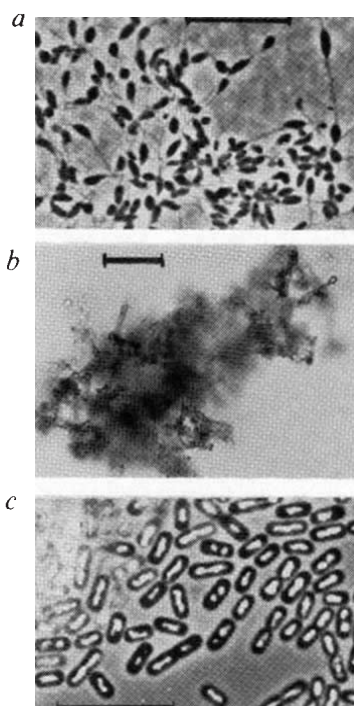


FIG. 2 Phase contrast photomicrographs of iron-oxidizing phototrophic bacteria. *a*, *Rhodomicrobium*-like isolate grown with succinate; *b*, *Rhodomicrobium*-like isolate grown with ferrous iron; *c*, gas vesicle-containing freshwater isolate grown with ferrous iron. Scale bars, 10 μ m.

did not change. No growth occurred in media without iron added.

These results show that ferrous iron oxidation requires only photosystem I in anoxygenic phototrophs, and that anaerobic ferrous iron oxidation was therefore possible before the evolution of oxygenic two-step photosynthesis. These findings, together with the recently discussed photochemical ferrous iron oxidation⁴⁻⁶, again question the assumption that the appearance of Fe(III) oxides in banded iron formations is always an indicator of the appearance of free oxygen. Anoxygenic phototrophic bacteria as isolated in the present study may have contributed to early banded iron formation at Archaean times before oxygen became significant as an oxidant. Such bacteria may still be of ecological significance as a means of regeneration of a widespread oxidant^{23,24} in anoxic low-sulphur littoral sediments. $\square$

Received 20 July 1992; accepted 17 February 1993.

1. Wood, P. M. in *Bacterial Energy Transduction* (ed. Anthony, C.) 183-230 (Academic, London, 1988).
2. Stumm, W. & Morgan, J. J. *Aquatic Chemistry* 2nd ed (Wiley-Interscience, New York, 1981).
3. Schwertmann, U. & Cornell, R. M. *Iron Oxides in the Laboratory* (VCH, Weinheim, 1991).
4. Cairns-Smith, A. G. *Nature* **276**, 807-808 (1978).
5. Braterman, P. S., Cairns-Smith, A. G. & Sloper, R. W. *Nature* **303**, 163-164 (1983).
6. François, L. M. *Nature* **320**, 352-354 (1986).
7. Hartman, H. in *Microbial Mats: Stromatolites* (eds Cohen, Y., Castenholz, R. W. & Halvorson, H. O.) 449-453 (Liss, New York, 1984).
8. Walker, J. C. G. *Nature* **329**, 710-711 (1987).
9. Cohen, Y. in *Microbial Mats* (eds Cohen, Y. & Rosenberg, E.) 22-36 (Am. Soc. Microbiol., Washington, 1990).
10. Schopf, J. W. *A. Rev. Earth Planet. Sci.* **3**, 213-249 (1975).
11. Schidlowski, M. in *The Early History of the Earth* (ed. Windley, B. F.) 525-535 (Wiley, London, 1976).
12. Gole, M. J. & Klein, C. *J. Geol.* **89**, 169-183 (1981).
13. Beukes, N. J. & Klein, C. in *The Proterozoic Biosphere* (eds Schopf, J. W. & Klein, C.) 147-158 (Cambridge Univ. Press, Cambridge, 1992).
14. Kasting, J. F., Holland, H. D. & Kump, L. R. in *The Proterozoic Biosphere* (eds Schopf, J. W. & Klein, C.) 159-163 (Cambridge Univ. Press, Cambridge, 1992).
15. Garrels, R. M. & Christ, C. L. *Solutions, Minerals and Equilibria* (Harper & Row, New York, 1965).
16. Mackenzie, F. T. & Wollast, R. in *Global Chemical Cycles and their Alterations by Man* (ed. Stumm, W.) 45-59 (Dahlem Konferenzen, Berlin, 1977).
17. Jørgensen, B. B. in *Microbial Geochemistry* (ed. Krumbein, W. E.) 91-124 (Blackwell, Oxford, 1983).
18. Dutton, P. L. & Prince, R. C. in *The Photosynthetic Bacteria* (eds Clayton, R. K. & Sistrom, W. R.) 525-570 (Plenum, New York, 1978).

19. Widdel, F. & Bak, F. in *The Prokaryotes* Vol. 4 (eds Balows, A., Trüper, H. G., Dworkin, M., Harder, W. & Schleifer, K.-H.) 3352-3378 (Springer, New York, 1992).
20. Pfennig, N. *Int. J. Syst. Bact.* **28**, 283-288 (1978).
21. Cohen, Y., Jørgensen, B. B., Revsbech, N. P. & Poplawski, R. *Appl. Environ. Microbiol.* **51**, 398-407 (1986).
22. Pfennig, N. in *Bergey's Manual of Systematic Bacteriology* Vol. 3 (eds Staley, J. T., Bryant, M. P., Pfennig, N. & Holt, J. G.) 1650-1651 (Williams & Wilkins, Baltimore, 1989).
23. Lovley, D. R. *Microbiol. Rev.* **55**, 259-287 (1991).
24. Nealson, K. H. & Myers, C. R. *Appl. Environ. Microbiol.* **58**, 439-443 (1992).
25. Stookey, L. L. *Analyt. Chem.* **42**, 779-781 (1970).
26. Lovley, D. R. & Philipps, E. J. P. *Appl. Environ. Microbiol.* **51**, 683-689 (1986).
27. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).

Kinship affects morphogenesis in cannibalistic salamanders

David W. Pfennig* & James P. Collins†

* Section of Neurobiology and Behavior, Mudd Hall, Cornell University, Ithaca, New York 14853, USA

† Department of Zoology, Arizona State University, Tempe, Arizona 85287, USA

INCLUSIVE fitness theory predicts that organisms can often increase their fitness by helping relatives¹. Indeed, many animals modify their behaviour towards kin in a fashion consistent with theory²⁻⁴. Morphogenesis may also be sensitive to kinship environment, especially in species that facultatively produce distinct morphs that differ in their ability to harm relatives, such as those that produce alternative cannibalistic and non-cannibalistic phenotypes⁵⁻⁹. We tested this hypothesis by examining whether consanguinity affected the probability that structurally distinctive cannibal morphs^{5,10} would develop in larval Arizona tiger salamanders (*Ambystoma tigrinum nebulosum*). We report here that when tiger salamander larvae are reared in mixed-brood groups they are significantly more likely to develop the cannibal morphology and at an earlier age than siblings reared in pure-sibship groups. In general, morphogenesis may be responsive to kinship in any species that facultatively develops structures that can be used against conspecifics as weaponry.

Tiger salamander larvae occur in nature as two alternative morphotypes^{5,10}: a 'typical' morph that feeds mostly on invertebrate prey, and a larger, physically distinctive 'cannibal' morph that has specialized oral structures to facilitate the ingestion of conspecifics¹⁰⁻¹². Cannibals are induced facultatively by high densities of conspecifics^{13,14}.

The potential exists in nature for cannibal-morph larvae to eat relatives. Stomach content analyses of wild-caught cannibals indicate that they often consume conspecifics of their own size and smaller^{11,15}. The occurrence of these two environmentally induced morphs enabled us to test whether kinship influences morphogenesis of alternative morphs that differ in their ability to harm relatives.

We randomly assigned similarly sized, two-week-old larvae from eight different sibships (three of which were cousins) to three different treatment categories: larvae were reared in groups of 16 with (1) siblings only, (2) equal numbers of siblings and nonsiblings (some were cousins, others were non-kin), or (3) one sibling and two larvae from each of the seven other sibships (Fig. 1). This manipulation mimicked a natural situation. Larvae are surrounded solely or primarily by siblings in ponds in which only a few females oviposit, whereas larvae are surrounded primarily by non-relatives in ponds where numerous females oviposit (unpublished observation). Inclusive fitness theory predicts that larvae in the former setting should be less likely than those in the latter setting to develop into cannibals.

Cannibal morphs were indeed significantly more likely to develop in mixed-brood groups than in pure-sibship groups (Table 1). No more than one cannibal was produced per aquarium. In aquaria containing relatively few larvae, cannibals

have the effect of inhibiting other larvae from developing into cannibal morphs¹³, either through chemical cues or by diminishing the food supply, thereby making it less profitable for another larva to develop into a cannibal. Thirty-one of 77 (40%) pure-sibship tanks produced a cannibal morph, as opposed to 33 of 40 (83%) two-sibship tanks, and 34 of 39 (87%) eight-sibship tanks. After we had statistically controlled for differences among sibships in the probability of producing a cannibal morph, we found that larvae in both the two- and eight-sibship treatments were significantly more likely to develop the cannibal morphotype than were their siblings reared in pure-sibship treatments (Table 1). Thus, increased consanguinity in the larval environment decreases a larva's probability of becoming a cannibal. Even small differences in kinship environment affected morphogenesis. The difference between observed and expected probabilities of an individual becoming a cannibal morph was significantly greater in the eight-sibship treatment (mean $\pm$ s.d. = 0.38 ± 0.34 ; $N = 39$) than in the two-sibship treatment (0.26 ± 0.35 ; $N = 40$; $P < 0.0001$, two-tailed Mann-Whitney test).

Presumably, a larva that becomes a cannibal later in life threatens its siblings less than one that becomes a cannibal early

in life. Thus, in addition to influencing the probability of a larva becoming a cannibal morph, kinship environment might affect the timing of the cannibal's initial expression.

The cannibal morphology was expressed significantly earlier in the eight-sibship treatment than in the one-sibship treatment (Fig. 2). Indeed, the mean time at which the cannibal morphotype developed in the eight-sibship treatment was earlier than in any of the eight sibships when they were reared alone. Thus, increased consanguinity in the larval environment also increased the age at which a larva became a cannibal. Interestingly, the cannibal morphotype developed significantly earlier in the two-sibship treatment containing non-kin and siblings than in the two-sibship treatment containing cousins and siblings (Fig. 2). This again illustrates that small differences in kinship environment influenced morphogenesis.

It may be argued that these differences between pure and mixed-sibship treatments were not due to differences in kinship

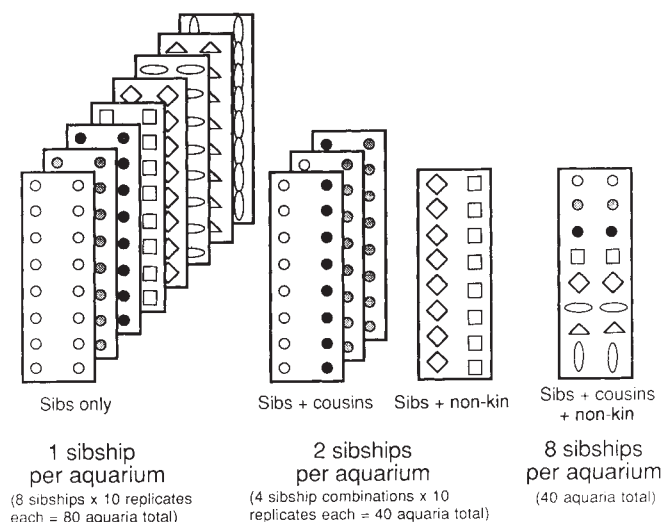


FIG. 1 Diagrammatic representation of the experimental design used to assess effects on morphogenesis of kinship. Each rectangle represents an aquarium containing 16 larvae, which are indicated by different geometric shapes (similar shapes shaded identically, siblings; similar shapes shaded differently, cousins; different shapes, non-kin). Study animals: we studied larvae from eight sibships whose parents or grandparents were caught in the White Mountains of Arizona, USA from five different ponds: Lower Cottonwood (LC), Snow (S), Old Apache (OA), South (So), and Wildcat Point (WP). Three sibships were captive-bred offspring of pairs collected from OA, So, and WP. Two sibships were captive-bred offspring of pairs collected from S, but were probably unrelated to each other owing to the large number of adults breeding in this pond. The other three sibships were offspring of captive-bred animals whose grandparents had been collected from LC and whose mothers were sisters. Thus, members of these three sibships were first cousins. Treatment groups: 160, 22-litre aquaria were each filled with 16 litres dechlorinated tap water. Two-week-old larvae (16 larvae per aquarium; tankmates differed in snout-vent length by no more than 5%) were distributed randomly¹³ among three treatment groups: (1) one sibship per aquarium; (2) two equally represented sibships per aquarium; or (3) eight sibships per aquarium. The range of within-group relatedness across treatments was within the range of estimates from natural ponds based on observations of numbers of breeding adults per pond. The rearing density also was within the range of larval densities in the field¹⁴. We followed the rearing procedures outlined in ref. 13. Morph assignments: the experiment was ended at seven weeks because cannibals are unlikely to develop after this time¹³. Tanks were coded by number. Without knowledge of which treatment group was in an aquarium, two persons independently scored each animal as being of the typical or cannibal morphotype using criteria described in refs 5 and 10. Morph assignments were always unambiguous.

TABLE 1 Effect of kinship environment on development of the cannibal morphology

Sibship(s)*	Number of replicates	Probability of producing a cannibal morph (number of aquaria in which a cannibal developed)		P
		Observed	Expected	
One sibship				
LC-16	10	0.20 (2)	—	
LC-21	9†	0.89 (8)	—	
LC-23	10	0.70 (7)	—	
S-4	8†	0.13 (1)	—	
S-5	10	0.30 (3)	—	
OA	10	0.40 (4)	—	
WP	10	0.20 (2)	—	
So	10	0.40 (4)	—	
Two sibships				
LC-16, LC-21	10	1.00 (10)	0.704 (7.04)	
LC-16, LC-23	10	0.70 (7)	0.819 (8.19)	
LC-21, LC-23	10	1.00 (10)	0.513 (5.13)	
S-4, S-5	10	0.60 (6)	0.222 (2.22)	
Total		(33)	(22.58)	0.0008
Eight sibships				
All sibships	39†	0.87 (34)	0.490 (19.11)	0.0001

Derivation of expected probability of producing a cannibal morph: the expected probability that a larva in each mixed-sibship tank would become a cannibal morph was calculated from the observed probabilities in the constituent pure-sibship groups. The observed probability that any given larva from a sibship would become a cannibal morph when reared with only siblings was $O_j = 1 - \sqrt[N_j]{1 - p_j}$, where p_j is the observed probability in Table 1 of sibship j producing a cannibal morph (for example, $O_{LC-16} = 1 - (1 - 0.2)^{1/16} = 0.014$). The expected probability of a cannibal morph developing in each mixed-sibship group was $E = 1 - \prod_{j=1}^S (1 - O_j)^{N_j}$, where O_j is observed probability that a larva from sibship j would become a cannibal morph when reared with siblings only, S is the number of sibships per tank (2 or 8), and N_j is the number of larvae from sibship j per tank (8 or 2). Because no more than one cannibal was produced per tank, we derived the expected probability that a cannibal morph would develop in S sibships (each containing N larvae) by subtracting from one the expected probability that those S sibships would not produce a cannibal morph ($(1 - O_1)^{N_1} \times (1 - O_2)^{N_2} \times \dots \times (1 - O_S)^{N_S}$). Sibships OA, WP and So were not used in the calculation of the expected probabilities for the two-sibship treatment because they were not represented in this treatment. Statistics: for each mixed-sibship treatment (that is, two and eight sibships per tank), the observed and expected number of aquaria in which a cannibal morph developed were compared with a one-sample χ^2 test. The four two-sibship combinations were pooled for analysis because they did not differ significantly in the difference between observed and expected probabilities ($H_{adj} = 6.78$; d.f. = 3; $P = 0.0792$; Kruskal-Wallis test).

* See Fig. 1 legend for notation used for study animals.

† Four tanks were excluded from analysis because all animals died from disease; mortality was low in the other tanks.

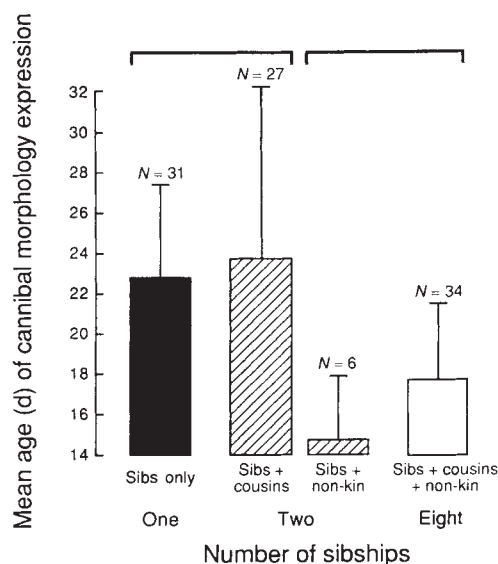


FIG. 2 Effect of kinship environment on larval age when the cannibal morph was initially expressed. Treatment group (number of sibships) had a highly significant effect on the time when a larva first developed the cannibal morphology ($H_{adj}=24.98$; d.f. = 3; $P < 0.0001$; Kruskal-Wallis test). Groups under different brackets were significantly different ($P < 0.005$), whereas groups under the same bracket were not (Kruskal-Wallis multiple comparisons test²⁷). Numbers above bars indicate the number of cannibal morphs produced per group.

per se. Instead, siblings may have been more similar in size than non-siblings, as in certain frog tadpoles¹⁶. If so, the benefits of becoming a cannibal might have been greater in mixed-sibship groups, because size disparities may facilitate cannibalism.

Four lines of evidence argue against this hypothesis. First, all larvae were similar in size initially, and individuals from different broods maintained similar growth trajectories, as evidenced by a lack of significant variation in snout-vent length (SVL) of randomly selected larvae from different sibships at the end of the experiment (typicals: $F_{4,36} = 0.71$, $P = 0.59$, $N = 40$ larvae from five sibships; cannibals: $F_{6,21} = 1.16$, $P = 0.36$, $N = 27$ larvae from seven sibships). Second, the size-variation hypothesis cannot explain why cannibals developed earlier in treatments containing siblings and non-kin than in those containing siblings and cousins (Fig. 2). If anything, cousins differed more in size (mean difference in SVL between cousins was 5.00 ± 5.71 mm; $N = 84$ pairwise combinations) than did non-kin (mean difference in SVL between non-kin was 4.41 ± 4.27 mm; $N = 224$ pairwise combinations). Third, at the end of the experiment, the proportion of cannibals produced in a brood was not positively correlated with the coefficient of variation in SVL of individual larvae from that sibship ($r = -0.20$, $P = 0.70$, $N = 6$ sibships, 105 larvae). Fourth, of 20 four-week-old typicals from sibships LC-16 and LC-23 (mean SVL, 25.85 ± 0.26 mm) that were housed individually with 15 smaller siblings (mean SVL, 18.21 ± 0.14 mm), none developed the cannibal morphology, even after consuming their siblings. By contrast, individuals from these sibships that were housed with equal-sized siblings and maintained under identical conditions occasionally developed into cannibals (Table 1). Indeed, the benefits of becoming a cannibal may be greater in groups of similar, not

dissimilar, size, because the cannibal phenotype enables salamanders to eat conspecifics of their own size¹². In sum, the size-variation hypothesis does not explain why larvae reared in mixed-brood groups were significantly more likely to develop the cannibal morphology than were their siblings reared in pure-sibship groups.

We hypothesize that expression of the cannibal morph is influenced by sibship-specific olfactory signals, which these larvae use to cannibalize more distant relatives in preference to close kin (D.W.P., P. Sherman and J.P.C., manuscript submitted). Sibship-specific chemical cues seem to cause certain frog tadpoles to grow faster in water conditioned by kin as opposed to non-kin¹⁷. Indeed, tadpoles of several anuran species grow larger when reared in pure-sibship groups than when reared in mixed-sibship groups¹⁷⁻²⁰.

Various studies have demonstrated that kinship environment can influence an organism's behaviour²¹ and that social environment can influence morphogenesis²². It is unknown, however, whether kinship affects morphogenesis in other species that produce alternative cannibalistic and non-cannibalistic morphs⁶⁻⁹ or 'fighter' and 'non-fighter' morphs^{23,24}, but there are some intriguing possibilities. For instance, different species of similarly sized parasitic wasps may attack the same caterpillar. Larvae of species in which each female typically lays a single egg per host tend to develop formidable fighter mandibles, whereas larvae of species in which each female typically lays many eggs per host (and whose larvae therefore share their host with many siblings) lack these structures²⁵. Thus, although lethal rivalry can arise even among close relatives²⁴⁻²⁶, kin-mediated morphogenesis may provide a general mechanism to reduce this conflict. □

Received 28 October 1992; accepted 4 February 1993.

- Hamilton, W. D. *J. theor. Biol.* **7**, 1-52 (1964).
- Sherman, P. W. *Behav. Ecol. Sociobiol.* **8**, 251-259 (1981).
- Emlen, S. T. & Wrege, P. H. *Behav. Ecol. Sociobiol.* **25**, 303-319 (1989).
- Packer, C., Gilbert, D. A., Pusey, A. E. & O'Brien, S. J. *Nature* **351**, 562-565 (1991).
- Powers, J. H. *Stud. Univ. Nebraska* **7**, 197-274 (1907).
- Armstrong, J. T. *Ecology* **45**, 361-365 (1964).
- Gilbert, J. J. *Science* **181**, 63-66 (1973).
- Pfennig, D. W. *Evolution* **46**, 1408-1420 (1992).
- Waddell, D. R. in *Cannibalism, Ecology and Evolution among Diverse Taxa* (eds Elgar, M. A. & Crespi, B. J.) 85-101 (Oxford Univ. Press, Oxford, 1992).
- Pedersen, S. C. *Amphibia-Reptilia* **12**, 1-14 (1991).
- Collins, J. P. & Holomuzki, J. R. *Can. J. Zool.* **62**, 168-174 (1984).
- Reilly, S. M., Lauder, G. V. & Collins, J. P. *Copeia* **1992**, 672-679 (1992).
- Collins, J. P. & Cheek, J. E. *Amer. Zool.* **23**, 77-84 (1983).
- Pfennig, D. W., Loeb, M. L. G. & Collins, J. P. *Oecologia* **88**, 161-166 (1991).
- Rose, F. L. & Armentrout, D. J. *anim. Ecol.* **45**, 713-729 (1976).
- Travis, J. *Evolution* **34**, 40-50 (1980).

- Shvarts, S. S. & Pyastolova, O. A. *Sov. J. Ecol.* **1**, 58-62 (1970).
- Jasienski, M. *Oecologia* **77**, 407-413 (1988).
- Smith, D. C. *Evolution* **44**, 1529-1541 (1990).
- Waldman, B. in *Kin Recognition* (ed. Hepper, P. G.) 162-219 (Cambridge Univ. Press, Cambridge, 1991).
- Hepper, P. G. *Kin Recognition* (Cambridge Univ. Press, Cambridge, 1991).
- West-Eberhard, M. J. *A. Rev. Ecol. Syst.* **20**, 249-278 (1989).
- Crespi, B. J. *Behav. Ecol. Sociobiol.* **23**, 93-104 (1988).
- Grbic, M., Ode, P. J. & Strand, M. R. *Nature* **360**, 254-256 (1992).
- Godfray, H. C. J. *Am. Nat.* **129**, 221-233 (1987).
- Parker, G. A., Mock, D. W. & Lamey, T. C. *Am. Nat.* **133**, 846-868 (1989).
- Conover, W. J. *Practical Nonparametric Statistics* (Wiley, New York, 1980).

ACKNOWLEDGEMENTS. We thank H. Reeve for deriving the expected probabilities in Table 1; H. Reeve, P. Sherman and U. Muller for comments on the manuscript; J. Roth and L. Nienaber for laboratory assistance; and the White Mountain Apache Indian tribe and the Arizona Game and Fish Department for collecting permits. This work was supported by an NSF postdoctoral fellowship to D.P. and by a grant from the NSF to J.C.

A metalloproteinase inhibitor domain in Alzheimer amyloid protein precursor

Kaoru Miyazaki, Miki Hasegawa, Kayano Funahashi & Makoto Umeda

Division of Cell Biology, Kihara Institute for Biological Research, Yokohama City University, 2-120-3 Nakamura-cho, Minami-ku, Yokohama 232, Japan

EXTRACELLULAR deposition of amyloid β -protein (β -AP), or A4 protein (M_r 4,000), is associated with Alzheimer's disease and with Down's syndrome (trisomy for chromosome 21)¹⁻³. The large membrane-bound precursor protein (APP) of β -AP is normally cleaved within the β -AP region by a putative proteinase (APP secretase) to release its extracellular portion; β -AP is produced by an alternative proteolytic processing⁴⁻⁶. Here we demonstrate that APP contains a proteinase inhibitor domain for the matrix metalloproteinase gelatinase A, which is located in the C-terminal glycosylated region of the secretory forms of APP. In addition, we show that the gelatinase has an APP secretase-like activity, which hydrolyses the Lys16-Leu17 bond in the β -AP sequence. Our results indicate that the proteinase inhibitor domain of APP and gelatinase A may be involved in the formation of β -AP.

To investigate how the extracellular proteolysis of APP is regulated, we used reverse zymography to analyse protein inhibitors of the matrix-degrading metalloproteinases (MMPs) secreted by different cultured cells. For the analysis, gelatinase A (72K gelatinase/type IV collagenase, or MMP2), the most common member of the MMPs⁷⁻⁹, was used as an indicator enzyme. Our survey revealed that many human cancer cell lines, including the EJ-1 bladder carcinoma cell line, secreted a novel gelatinase inhibitor of M_r 100K as well as TIMP (tissue inhibitor of metalloproteinases; M_r 28K), TIMP-2 (20K), and an unidentified 22K inhibitor (Fig. 1a). The 100K inhibitor purified from serum-free conditioned medium of EJ-1 cells ran as a major band at 100K and as a minor band of 90K on sodium dodecyl sulphate (SDS)-polyacrylamide gel electrophoresis (PAGE) under non-reducing conditions (110K and 95K under reducing conditions, respectively) (Fig. 1b, c). The 100K inhibitor has an

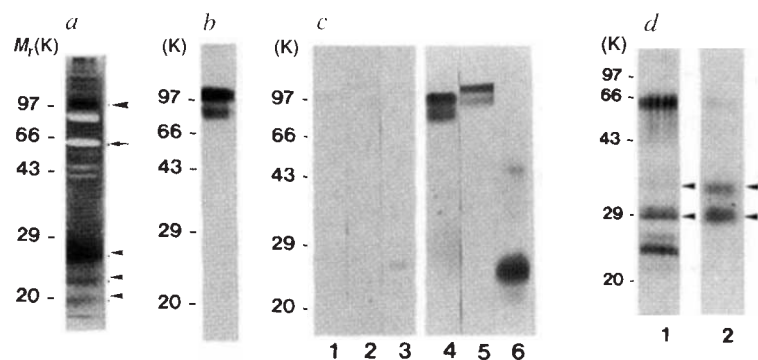
isoelectric point (pI) of 3.6 at 5 °C by isoelectric focusing. The 100K and 90K proteins both showed gelatinase inhibitor activity on reverse zymography before and after reduction with 2-mercaptoethanol (Fig. 1c); these activities are comparable to that of TIMP.

The purified 100K inhibitor has an N-terminal sequence Leu-Glu-Val-Pro-Thr-Asp-Gly-Asn-Ala-Gly-Leu- (Fig. 2a, b); this sequence is identical to that of APP without the signal sequence. Three principal APP forms having 695, 751 and 770 amino acids (APP695, APP751 and APP770, respectively) are known to be produced by alternative messenger RNA splicing from a single APP gene located on chromosome 21 (refs 4, 10-12). APP751 and APP770 contain a Kunitz-type serine-proteinase inhibitor (KPI) domain¹⁰⁻¹², and their secretory forms are identical to protease nexin-II (refs 13, 14). The 100K and 90K gelatinase inhibitors also inhibit trypsin and chymotrypsin, indicating that they are the secretory forms of either APP770 or APP751.

APP molecules can be divided into distinct regions (Fig. 2a; refs 4, 10-12). To localize the gelatinase inhibitor domain, we digested the purified protein with the endoproteinase Lys-C. The digest contained two active fragments of 28K and 31K, as analysed by reverse zymography (Fig. 1d). The main 28K fragment has an N-terminal sequence Val-Glu-Ser-Leu-Glu-Gln-Glu-Ala- (Fig. 2c); this sequence is located in the C-terminal glycosylated region of the secreted APP. Unlike the KPI domain, this region exists in all APP forms, including APP695, the brain-specific APP. There is no sequence homology between the gelatinase inhibitor region of APP and the known metalloproteinase inhibitors TIMP and TIMP-2.

The normal processing of APP by its secretase produces the ~100K secretory form and a residual 9K membrane-bound fragment by cleavage of the Lys 16-Leu 17 bond within the β -AP region^{6,15} (Fig. 2d). We therefore also tested the APP secretase activity of gelatinase A using a synthetic peptide consisting of residues 10-20 of β -AP (β -AP₁₀₋₂₀) as a substrate (Fig. 2d). We have previously shown that two active forms (57K and 41K) are produced from latent 64K progelatinase A by a combination of activating reagents such as *p*-aminophenyl mercuric acetate (APMA), SDS and stromelysin⁹. Treatment of progelatinase A free of TIMP-2 with 1 mM APMA at 37 °C for 1 h produced only the 57K form, whereas excessive treatment with APMA, or short treatment with APMA plus stromelysin, produced both the 57K and 41K forms. When the APMA/stromelysin-activated

FIG. 1 Electrophoretic analyses of 100K inhibitor secreted by EJ-1 cells. **a**, Reverse zymography of the conditioned medium of EJ-1 cells. Large arrowhead, 100K inhibitor; small arrowheads, TIMP (28K), unidentified 22K inhibitor and TIMP-2 (20K). The negatively stained bands are gelatinolytic activities secreted by the tumour cells: arrow, gelatinase A. **b**, Non-reducing SDS-PAGE of purified 100K inhibitor. **c**, SDS-PAGE (lanes 1-3) and reverse zymography (lanes 4-6) of 100K inhibitor (lanes 1, 2, 4 and 5) and human TIMP (lanes 3 and 6). The same amount (0.2 μ g per lane) of each protein was analysed under reducing (lanes 2 and 5) and non-reducing conditions (lanes 1, 3, 4 and 6). Note that in these amounts, both proteins are only faintly stained by Coomassie brilliant blue (CBB) on SDS-PAGE, but they showed strongly stained gelatin bands (inhibitor bands) on reverse zymography. **d**, SDS-PAGE (lane 1) and reverse zymography (lane 2) of the 100K inhibitor after treatment with endoproteinase Lys-C. Arrowheads indicate two active fragments of M_r 28K and 31K. **METHODS.** **a**, Conditioned medium was prepared from the confluent culture of EJ-1 cells incubated for 2 d in serum-free RPMI 1640 medium, concentrated 30-fold, and electrophoresed as described²². Reverse zymography of gelatinase inhibitors was on a polyacrylamide gel containing 0.1% SDS and 1 mg ml⁻¹ gelatin according to ref. 23, with modifications. After electrophoresis and subsequent renaturation^{8,22}, the gel was incubated at 37 °C for 18 h in 4 ml reaction mixture containing 50 mM Tris-HCl, pH 7.5, 10 mM CaCl₂, 1 μ g ml⁻¹ human progelatinase A and 1 mM *p*-aminophenyl mercuric acetate (APMA), followed by staining with CBB. Progelatinase A, which contained the TIMP-2-bound form (<85%), the TIMP-2-free form (<15%)



and small amounts of 57K and 41K active forms, was prepared from conditioned medium of the T98G human glioblastoma cell line by a method reported earlier⁹. **b**, About 1.1 mg 100K inhibitor was purified from 6 l EJ-1 conditioned medium by chromatography on reactive red-agarose (Sigma) and a Shodex QA-824 anion-exchange HPLC column (Showa Denko, Tokyo). Both columns were pre-equilibrated with 20 mM Tris-HCl, pH 7.5 in 0.005% Brij 35, and proteins were eluted with increasing NaCl. **d**, 100K inhibitor (200 μ g) was incubated with 0.3 μ g endoproteinase Lys-C (Sigma) in 60 μ l 50 mM Tris-HCl, pH 7.5, containing 2 M urea and 1 M NaCl, at 30 °C for 60 min, and a sample of the digest taken for electrophoresis.

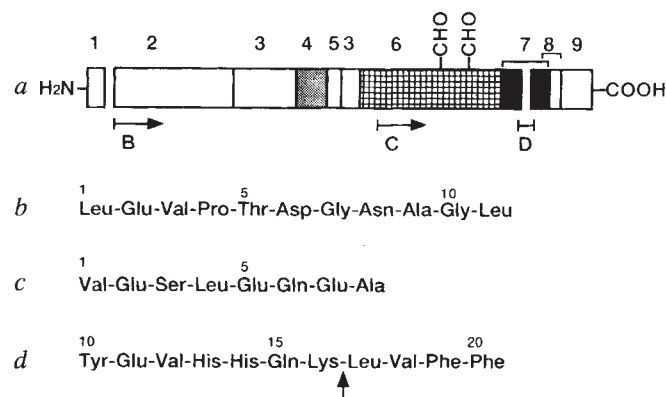


FIG. 2 Schematic illustration of APP770 structure and the partial amino-acid sequence of the 100K inhibitor. *a*, Domain structure of APP770: 1, signal sequence; 2, cysteine-rich region; 3, acidic region; 4, KPI domain; 5, 19-amino-acid insert; 6, glycosylated region (gelatinase inhibitor domain revealed in our study); 7, β -AP region; 8, transmembrane region; 9, cytoplasmic region. APP751 lacks region 5, and APP695 lacks regions 4 and 5. *b*, N-terminal amino-acid sequence of the purified 100K inhibitor. The sequence corresponds to residues 18–28 of the intact APP (arrow B in *a*). *c*, N-terminal amino-acid sequence for the 28K gelatinase inhibitor fragment. The sequence corresponds to residues 439–446 of the intact APP770 (arrow C in *a*). *d*, Partial β -AP sequence (β -AP_{10–20}) containing the normal proteolytic processing site of APP. The sequence corresponds to residues 681–691 of the intact APP770 (bar D in *a*). Arrow, cleavage site by gelatinase A.

METHODS. For protein sequencing, the purified 100K gelatinase inhibitor and its 28K active fragment (*d*) were run separately on SDS-PAGE, electroblotted from the SDS gels onto polyvinylidene difluoride membranes and then applied to an automated gas-phase protein sequencer (model 477A; Applied Biosystems).

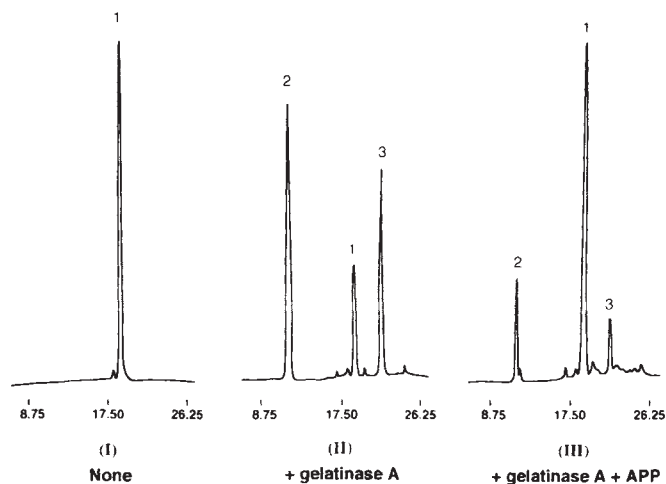


FIG. 3 β -AP peptide hydrolysis of gelatinase A and its inhibition by 100K APP. β -AP_{10–20} was analysed by reverse-phase HPLC before (I) and after incubation with the APMA/stromelysin-activated gelatinase A in the absence (II) or presence (III) of the 100K APP. Peak 1, β -AP_{10–20}; peak 2, β -AP_{10–16} (Tyr-Glu-Val-His-His-Gln-Lys); peak 3, β -AP_{17–20} (Leu-Val-Phe-Phe). Abscissa shows elution time in minutes.

METHODS. TIMP-2-free progelatinase A was separated from the TIMP-2-bound form by heparin-affinity chromatography according to ref. 24. The proenzyme (1 μ g) was activated by incubating with 1 mM APMA and 0.25 μ g rat stromelysin in 20 mM Tris-HCl, pH 7.5, containing 10 mM Ca^{2+} at 37 °C for 1 h. The activated enzyme was then incubated with 20 nmol (28 μ g) of β -AP_{10–20} (Fig. 2*d*) in the absence or presence of 5 μ g 100K APP in 50 μ l Tris-HCl/ Ca^{2+} buffer at 37 °C for 6 h, mixed with 0.5 ml 0.05% trifluoroacetic acid (TFA), and then applied to a Cosmosil 5C18 reverse-phase HPLC column (4.6 $\times$ 150 mm) (Nacalai Tesque, Kyoto). The loaded column was eluted with a linear gradient of 0–80% acetonitrile in 15 ml 0.05% TFA at a flow rate of 0.5 ml min⁻¹. Amino-acid sequences of the indicated peptides were analysed with the protein sequencer. Rat stromelysin was purified in a proenzyme form as before and activated by APMA⁹. Stromelysin barely hydrolyses β -AP_{10–20} under these conditions.

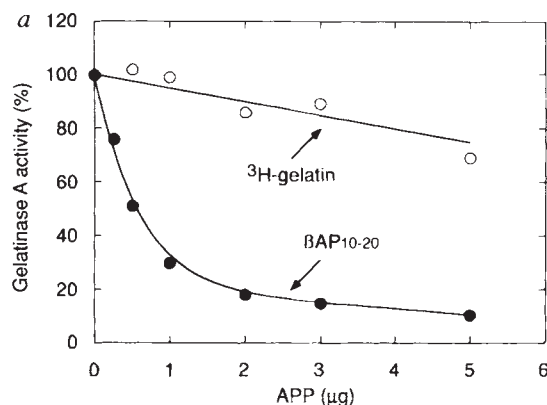
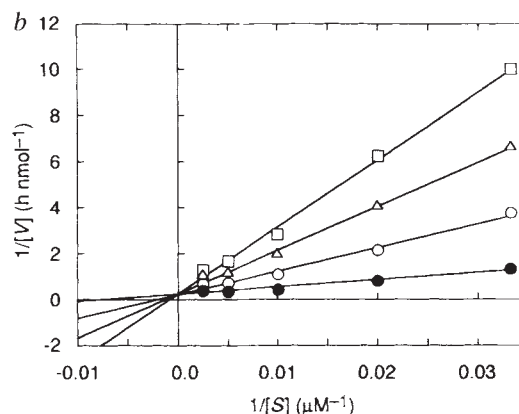


FIG. 4 Kinetic analysis of inhibition of gelatinase A activity by 100K APP. *a*, Effect of inhibitor concentration on β -AP_{10–20} (●) and ³H-gelatin (○) hydrolysis. The highest concentration of 100K APP (5 μ g per 50 μ l) is equivalent to ~900 nM. *b*, Lineweaver-Burk plot (1/*V* versus 1/[*S*] plot) for β -AP_{10–20} hydrolysis in the absence (●) or presence of 91 nM (○), 182 nM (△), or 364 nM (□) 100K APP.

METHODS. *a*, β -AP_{10–20} hydrolysis of gelatinase A was assayed as described for Fig. 3, except that 0.5 μ g progelatinase A activated with APMA plus 0.125 μ g stromelysin was incubated for 30 min with 0.2 mM β -AP_{10–20} in the presence of 0–5.0 μ g (0–910 nM) of 100K APP in 50 μ l. β -AP_{10–20} hydrolysis was determined from the area of the β -AP_{10–16} peak (Fig. 3, II) in the HPLC profiles. The ³H-gelatin-hydrolysing activity was assayed accord-



ing to ref. 25 using ³H-labelled human placental type IV collagen (NEN; 0.31 mCi mg⁻¹), previously denatured at 60 °C for 10 min, as a substrate. Activated gelatinase (0.5 μ g) was incubated with 20 nCi of ³H-gelatin in 50 μ l reaction mixture at 37 °C for 6 h. The ³H-gelatin hydrolysis by stromelysin, which occupied ~10% of the total activity, was subtracted. *b*, The activated progelatinase was incubated with 30–400 μ M β -AP_{10–20} in the presence or absence of 100K APP. The *K_i* value (40 nM) of 100K APP for β -AP_{10–20} hydrolysis was calculated from the slope of the Dixon plot (1/*V* versus 1/[*S*] plot) of the data in *a* and the *K_m* (130 μ M) and *V_{max}* (3.9 nmol h⁻¹) values obtained from *b*, using the equation $1/V = 1/V_{max} + K_m(1 + (I/K_i))/V_{max}[S]$.

gelatinase A was incubated with β -AP₁₀₋₂₀, the peptide was effectively hydrolysed into two fragments (Fig. 3, I and II). These proteolytic fragments correspond to β -AP₁₀₋₁₆ (Tyr-Glu-Val-His-His-Gln-Lys) and β -AP₁₇₋₂₀ (Leu-Val-Phe-Phe), indicating that gelatinase A hydrolyses the Lys 16-Leu 17 bond, which accords with the substrate specificity of APP secretase (Fig. 2d). The isolated 41K enzyme hydrolyses both β -AP₁₀₋₂₀ and gelatin, whereas the 57K enzyme hydrolyses only gelatin (data not shown).

The 100K APP effectively inhibits β -AP₁₀₋₂₀ hydrolysis of activated gelatinase A, although ³H-gelatin hydrolysis is inhibited far less (Fig. 3, III, and Fig. 4a). The APP concentration giving 50 per cent inhibition (IC₅₀) was ~90 nM for β -AP₁₀₋₂₀ hydrolysis and ~1.5 μ M for gelatinolysis. Kinetic analysis at various concentrations of β -AP₁₀₋₂₀ and inhibitor indicated that the K_m of gelatinase A for β -AP₁₀₋₂₀ is ~130 μ M, and that 100K APP inhibits β -AP₁₀₋₂₀ hydrolysis in a competitive manner (Fig. 4b). From the data shown in Fig. 4a and b, the inhibition constant (K_i) of 100K APP for β -AP₁₀₋₂₀ hydrolysis is estimated as ~40 nM.

Gelatinase A exists in the plasma membrane of cells in lines¹⁶ and its proenzyme is activated to the 57K and 41K forms by a plasma membrane fraction¹⁷. We have detected both gelatinase A and soluble APP in high amounts in conditioned media of human cancer cell lines such as glioblastoma, neuroblastoma and EJ-1 (Fig. 1a), and in human cerebrospinal fluids. Over-expression of the APP gene in neuronally differentiated mouse embryonal carcinoma cells is known to cause aberrant processing of APP and degenerative death of these cells¹⁸; also, both the secretory APP and β -AP forms are produced in culture and in normal individuals¹⁹⁻²¹. Our results suggest two possible functions for gelatinase A: it may produce secretory APP

on the plasma membrane and it may degrade soluble β -AP in the extracellular matrix. Both actions would prevent β -AP accumulation in the extracellular matrix. When APP is over-produced under certain conditions, however, it may inhibit these activities of gelatinase A, favouring the extracellular deposition of β -AP. □

Received 30 November 1992; accepted 18 February 1993.

1. Glenner, G. G. & Wong, C. W. *Biochem. biophys. Res. Commun.* **122**, 1131-1159 (1984).
2. Masters, C. L. et al. *Proc. natn. Acad. Sci. U.S.A.* **82**, 4245-4349 (1985).
3. Yankner, B. A. et al. *Science* **245**, 417-420 (1989).
4. Kang, J. et al. *Nature* **325**, 733-736 (1987).
5. Sisodia, S. S., Koo, E. H., Beyreuther, K., Unterbeck, A. & Price, D. L. *Science* **248**, 492-495 (1990).
6. Esch, F. S. et al. *Science* **248**, 1122-1124 (1990).
7. Collier, I. E. et al. *J. biol. Chem.* **263**, 6579-6587 (1988).
8. Yasumitsu, H., Miyazaki, K., Umenishi, F., Koshikawa, N. & Umeda, M. *J. Biochem. (Tokyo)*, **111**, 74-80 (1992).
9. Miyazaki, K. et al. *Biochem. biophys. Res. Commun.* **185**, 852-859 (1992).
10. Ponte, P. et al. *Nature* **331**, 525-527 (1988).
11. Tanzi, R. E. *Nature* **331**, 528-530 (1988).
12. Kitaguchi, N., Takahashi, Y., Tokushima, Y., Shiojiri, S. & Ito, H. *Nature* **331**, 530-532 (1988).
13. Oltersdorf, T. et al. *Nature* **341**, 144-147 (1989).
14. Van Nostrand, W. E. et al. *Nature* **341**, 546-549 (1989).
15. Anderson, J. P. et al. *Neurosci. Lett.* **128**, 126-128 (1991).
16. Zucker, S. et al. *Int. J. Cancer* **45**, 1137-1142 (1990).
17. Murphy, G. et al. *Biochem. J.* **283**, 637-641 (1992).
18. Yoshikawa, K., Aizawa, T. & Hayashi, Y. *Nature* **359**, 64-67 (1992).
19. Haass, C. et al. *Nature* **359**, 322-325 (1992).
20. Seubert, P. et al. *Nature* **359**, 325-327 (1992).
21. Shoji, M. et al. *Science* **258**, 126-129 (1992).
22. Koshikawa, N., Yasumitsu, H., Umeda, M. & Miyazaki, K. *Cancer Res.* **52**, 5046-5053 (1992).
23. Herron, G. S., Banda, M. J., Clark, E. J., Gavrilovic, J. & Werb, Z. *J. biol. Chem.* **261**, 2814-2818 (1986).
24. Kolkenbrock, H., Orgel, D., Hecker-Kia, A., Noack, W. & Ulbrich, N. *Eur. J. Biochem.* **198**, 775-781 (1991).
25. Liotta, L. A. et al. *Nature* **284**, 67-68 (1980).

ACKNOWLEDGEMENTS. We thank H. Yasumitsu, Y. Suzuki, N. Koshikawa and Y. Numata for technical support and for discussion. This work was supported in part by a grant from the Oriental Yeast Co. (Tokyo) and the Special Coordination Funds of the Science and technology Agency of Japan.

Inhibition of vascular endothelial growth factor-induced angiogenesis suppresses tumour growth *in vivo*

K. Jin Kim, Bing Li, Jane Winer, Mark Armanini, Nancy Gillett, Heidi S. Phillips & Napoleone Ferrara*

Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, USA

THE development of new blood vessels (angiogenesis) is required for many physiological processes including embryogenesis, wound healing and corpus luteum formation^{1,2}. Blood vessel neof ormation is also important in the pathogenesis of many disorders¹⁻⁵, particularly rapid growth and metastasis of solid tumours³⁻⁵. There are several potential mediators of tumour angiogenesis, including basic and acidic fibroblast growth factors, tumour necrosis factor- α and transforming factors- α and - β ^{1,2}. But it is unclear whether any of these agents actually mediates angiogenesis and tumour growth *in vivo*. Vascular endothelial growth factor (VEGF) is an endothelial cell-specific mitogen and an angiogenesis inducer released by a variety of tumour cells and expressed in human tumours *in situ*. To test whether VEGF may be a tumour angiogenesis factor *in vivo*, we injected human rhabdomyosarcoma, glioblastoma multiforme or leiomyosarcoma cell lines into nude mice. We report here that treatment with a monoclonal antibody specific for VEGF inhibited the growth of the tumours, but had no effect on the growth rate of the tumour cells *in vitro*. The density of vessels was decreased in the antibody-treated

tumours. These findings demonstrate that inhibition of the action of an angiogenic factor spontaneously produced by tumour cells may suppress tumour growth *in vivo*.

VEGF has several attractive features as a mediator of normal and pathological angiogenesis. VEGF is an endothelial cell-specific mitogen and an angiogenesis inducer *in vivo*⁶⁻⁹. Its high-affinity binding sites are localized only on endothelial cells in tissue sections¹⁰. VEGF was also purified independently as a tumour-derived vascular permeability factor^{11,12}. By alternative splicing of messenger RNA, VEGF may exist in four different homodimeric molecular species, each monomer having 121, 165, 189 or 206 amino acids, respectively (VEGF₁₂₁, VEGF₁₆₅, VEGF₁₈₉, VEGF₂₀₆)¹³⁻¹⁵. VEGF₁₂₁ and VEGF₁₆₅ are soluble proteins, whereas VEGF₁₈₉ and VEGF₂₀₆ are bound to heparin-containing proteoglycans in the cell surface or in the basement membrane¹⁶. A temporal and spatial correlation exists between VEGF mRNA expression and physiological proliferation of blood vessels in the ovarian corpus luteum¹⁷ or in the developing brain¹⁸. Several observations suggest that VEGF may be potentially involved in tumour angiogenesis. A variety of transformed cell lines express the VEGF mRNA and secrete VEGF^{19,20}. Also, *in situ* hybridization studies demonstrate expression of VEGF mRNA at high level in various human tumours, including the highly vascularized glioblastoma multiforme and capillary haemangioma²¹⁻²⁴. Furthermore, expression of VEGF₁₆₅ or VEGF₁₂₁ confers on Chinese hamster ovary cells the ability to form tumours in nude mice²⁵.

Proof of the role of VEGF in tumour angiogenesis requires the demonstration that inhibition of VEGF action prevents tumour growth *in vivo*. The availability of specific monoclonal antibodies capable of blocking VEGF-induced angiogenesis *in vivo* and *in vitro*²⁶ allowed us to test the hypothesis directly. The human A673 rhabdomyosarcoma, G55 glioblastoma multiforme and SK-LMS-1 leiomyosarcoma cell lines are tumorigenic in nude mice, express the VEGF mRNA (ref. 20 and unpublished

* To whom correspondence should be addressed.

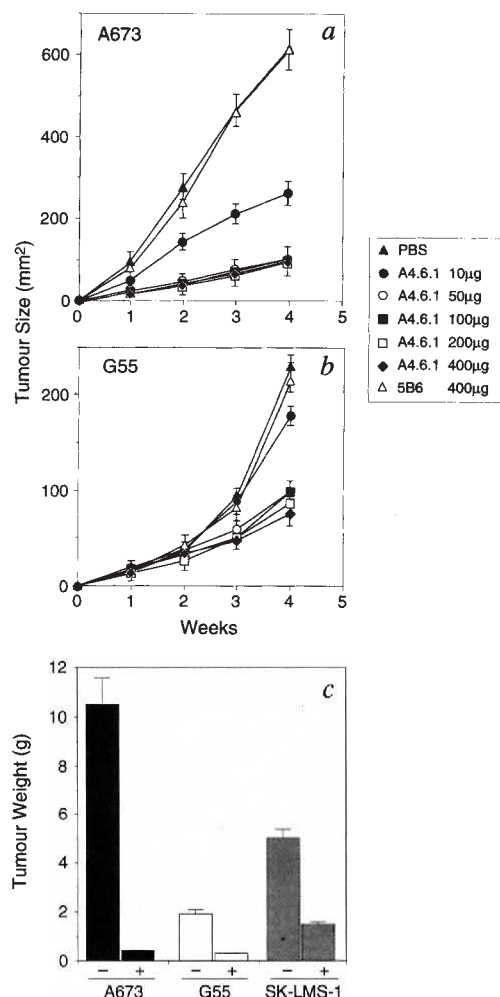


FIG. 1 Effects of anti-VEGF antibody on tumour size (a, b) and weight (c). A673 (CRL 1598) and SK-LMS-1 (HTB 88) cell lines were from the ATCC. A series of cell lines were established from glioblastoma multiforme specimens³⁰. The G55 cell line was reproducibly tumorigenic and secreted VEGF. Cells were cultured in the presence of DMEM/F12 supplemented with 10% FBS, 2 mM glutamine and antibiotics and passaged weekly at a split ratio of 1:20. For injection into nude mice, cells (passages 3–5) were dissociated from stock plates in the presence of 125 mM NaCl, 5 mM KCl, 50 mM HEPES, 5 mM glucose, 1 mM EDTA, pH 7.4, washed once and then resuspended in PBS at the desired density. Female Beige nude/xid mice (6–10 week old) (Charles River, Wilmington, DE) were injected subcutaneously with 1×10^6 cells in the dorsal area in a volume of 0.1 ml. The neutralizing anti-VEGF antibody A.4.6.1 (ref. 26) was injected intraperitoneally in a volume of 0.2 ml twice weekly at various doses. As controls, an antibody of the same isotype (IgG₁), directed against the gp120 protein³¹ (antibody 5B6) or PBS were used. Each group comprised 10 mice. In a, the antibody treatment was initiated immediately after cell injection. In b, it was initiated after 7 days. At the end of the experiment, animals were killed by CO₂ inhalation. Tumour sizes were determined by multiplying the width by the length. Values shown are means $\pm$ s.e.m. Statistical analysis was done by a two-sided unpaired *t*-test following a one-way analysis of variance. Comparison of tumour size of the PBS group with those of animals treated with antibody A.4.6.1 at doses of 50 µg or higher revealed *P* values < 0.0001 for both types of tumour. *P* values for the 10 µg dose were 0.003 for the A673 and 0.011 for the G55. Tumour sizes in the group treated with antibody 5B6 were not different from those injected with PBS alone. c. Weights of tumours derived from A673, G55 or SK-LMS-1 cells. Data shown reflect the response to 100 µg twice weekly of antibody A.4.6.1. Values are means $\pm$ s.e.m. Plus and minus signs denote the presence or absence of antibody treatment. A673 and G55 tumours were collected 4 weeks after tumour cell injection. Animals bearing tumours derived from SK-LMS-1 cells were killed after 10 weeks.

observations), and release VEGF in the culture medium (Table 1). The A673 cells secrete the highest amounts of VEGF, the SK-LMS-1 cells the lowest. Such cells may thus provide a model to test the hypothesis that VEGF is a paracrine mediator of tumour growth in an *in vivo* system.

We injected the tumour cells into nude mice. Animals then received various doses of an anti-VEGF monoclonal antibody, a control antibody or PBS twice weekly. Figures 1a and b illustrate the growth, as a function of time, of tumours derived from A673 or G55 cells. Although the control antibody had no appreciable effect, the anti-VEGF antibody was growth inhibitory to both types of tumour. But the magnitude of the response was greater in the rhabdomyosarcoma, the more rapidly proliferating and therefore the more angiogenesis-dependent tumour. As little as 10 µg twice weekly of the antibody resulted in a significant inhibition. In both tumours, a maximal effect was achieved with antibody doses of 50–100 µg. Growth inhibition was achieved not only when the anti-VEGF antibody administration was started immediately after tumour cell inoculation but also when such treatment was initiated a week later, when the

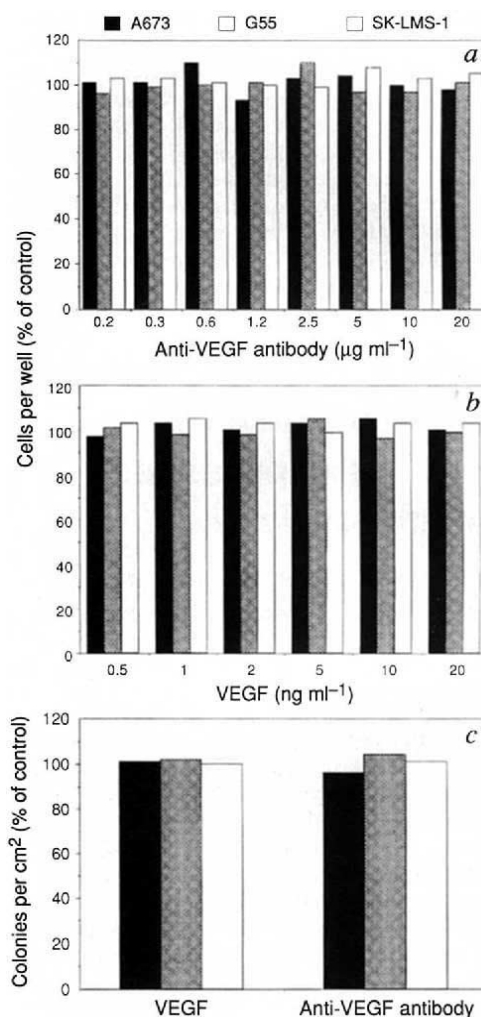


FIG. 2 Effects of VEGF or anti-VEGF antibody on the growth of tumour cell lines in plastic or in soft agar. For proliferation assays, tumour cells were seeded at the density of 7×10^3 per well in 12 multiwell plates in the presence of DMEM/F12 supplemented with 10% FBS, 2 mM glutamine and antibiotics: rhVEGF₁₆₅ or antibody A.4.6.1 were added at the indicated concentrations. After 5 days, triplicate wells were exposed to trypsin and cells were counted in a Coulter counter. Soft agar colony formation assay was done as described²⁹. rhVEGF₁₆₅ was added at the final concentration of 20 ng ml⁻¹; antibody A.4.6.1 was tested at the final concentration of 1 µg ml⁻¹.

tumours were already established (Fig. 1b). The SK-LMS-1 cells had the slowest growth rate and tumours could not be measured for several weeks. The anti-VEGF antibody also inhibited their growth. Figure 1c shows the weight of tumours derived from the three cell lines after treatment with 100 μ g twice weekly of anti-VEGF antibody in a representative experiment. The decreases in tumour weight were 96% for the A673, 80% for the G55 and 70% for the SK-LMS-1 cells. Such profound effects are consistent with inhibition of a fundamental mechanism of tumour growth. The lack of complete inhibition may be explained by the presence of additional angiogenic factors. For example, basic or acidic fibroblast growth factors are produced by glioma²⁷ and sarcoma²⁸ cell lines.

To verify that such *in vivo* effects do not result from inhibition of autocrine actions of VEGF or from direct cytotoxicity of the antibody, we tested the anti-VEGF antibody or VEGF (Fig. 2) for their ability to affect the *in vitro* growth of tumour cell lines, both in plastic²⁵ and in soft agar²⁹. Neither the antibody nor VEGF had any effect on the growth rate or colony formation in any of the cell lines tested. These findings agree with previous studies that showed VEGF to be an endothelial cell-specific mitogen⁶.

Microscopic examination of tumours and other tissues from animals injected with A673 or G55 cells did not reveal marked differences between control and anti-VEGF antibody-treated groups in several histological parameters. Multifocal coalescing foci of necrosis interspersed among masses of tumour cells were present in both rhabdomyosarcoma and glioblastoma multiforme sections. Metastases to lungs or liver were very rare. Inflammatory cell infiltrates were minimal and oedema was virtually absent in all tumour sections examined. But the density of vascular elements was decreased in the anti-VEGF antibody-treated tumours when compared with controls. This was confirmed by immunocytochemistry with an antiserum directed against factor VIII-related antigen (Fig. 3a-c). This decrease in vascular density was more prominent in sections from the rhabdomyosarcoma than from the glioblastoma. This is consistent with the more dramatic response of the rhabdomyosarcoma to the anti-VEGF antibody.

TABLE 1 Concentrations of VEGF by ELISA in the conditioned medium of tumour cell lines

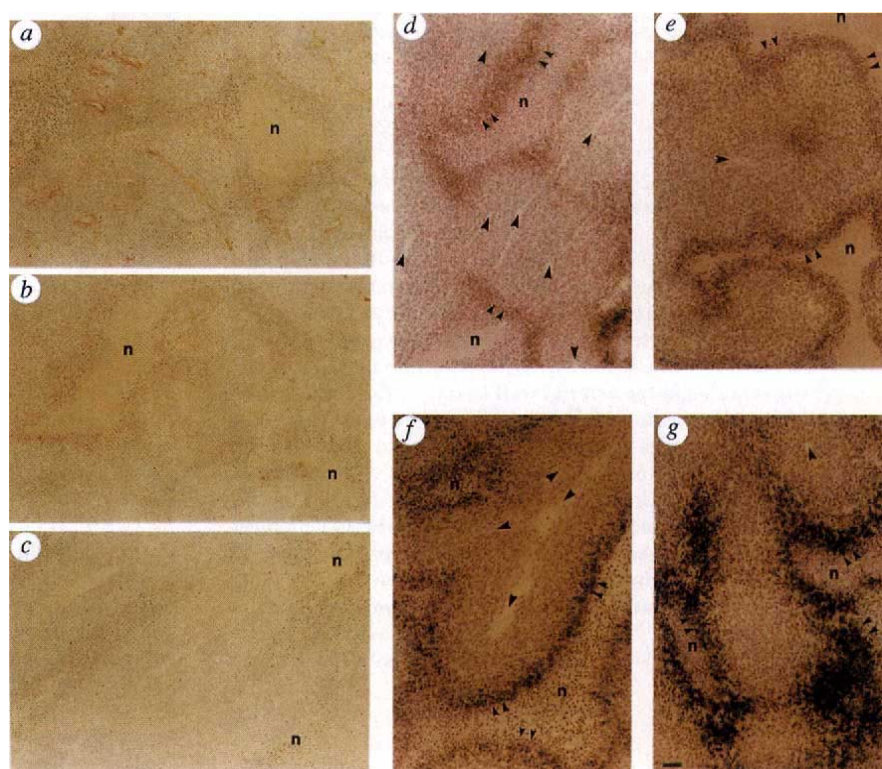
	48 h	96 h	144 h
A673	5	45	85
G55	2	18	41
SK-LMS-1	<0.2	2	14

Cells were cultured in 10-cm tissue culture dishes in the presence of Dulbecco's modified Eagle's medium (DMEM)/F12 supplemented with 10% fetal bovine serum, 2 mM glutamine and antibiotics. Subconfluent cultures were washed and then incubated in serum-free medium. At the indicated time points, aliquots of medium were collected and subjected to enzyme-linked immunosorbent assay (ELISA). The VEGF ELISA was done as described¹⁶. Values are in ng ml⁻¹. The limit of sensitivity of the assay was 0.2 ng ml⁻¹.

To elucidate the expression pattern of VEGF mRNA, we did *in situ* hybridization with a probe specific for VEGF on tumours derived from A673 or G55 cells. We found similar patterns in the two types of tumour (Fig. 3d-g). In tumours from untreated animals, VEGF expression was not uniform but was at the highest level in clusters of cells at the periphery of cell cords surrounded by necrotic areas. Numerous blood vessels were identified within these tumour cell cords (Fig. 3a, d, f). Remarkably, this pattern is very similar to that recently described in glioblastoma multiforme specimens *in situ*^{22,23}. This is consistent with the hypothesis that hypoxaemia triggers expression and release of VEGF in focal areas of the tumour^{22,23}. This would generate a spatial gradient of endothelial cell proliferation and angiogenesis toward ischaemic areas, permitting cell proliferation and tumour expansion. In the anti-VEGF antibody-treated groups, VEGF expression was also highest in tumour cells facing necrotic areas. But blood vessel lumina were rarely observed (Fig. 3b, e, g). These findings strongly support the hypothesis that inhibition of angiogenesis is the event limiting tumour growth in the anti-VEGF antibody-treated animals.

The probe that we used does not discriminate between the

FIG. 3 Immunocytochemical stain for factor VIII-related antigen (a-c) and *in situ* hybridization with a probe specific for VEGF (d-g). a, c, A673 tumours from untreated or b, antibody-treated animals. In c, normal rabbit serum was added instead of anti-factor VIII antiserum. Tissues were fixed in the presence of 10% neutral buffered phosphate formalin and then paraffin-embedded. Immunocytochemistry was done with the avidin-biotin-peroxidase kit with diaminobenzidine as the chromogen, according to the instructions of the manufacturer (Vector Labs, Burlingame, CA). The primary antiserum (rabbit anti-human factor VIII; Dako, Carpinteria, CA) or normal rabbit serum was added at the dilution of 1:100 for 30 min at room temperature. d, e, Brightfield images of A673 or f, g, G55 tumour sections hybridized with an antisense RNA probe specific for human VEGF. d, f, Untreated. e, g, Treated with anti-VEGF antibody. Individual arrows point to blood vessel lumina. Double arrows point to VEGF-expressing cells. *In situ* hybridization was done on unfixed frozen sections using a 1 kilobase probe representing residues 924-1,920 of a human VEGF complementary DNA, as described³². No appreciable hybridization was observed in sections incubated with the sense probe. Tumours used for immunocytochemistry or *in situ* hybridization were collected 4 weeks after cell inoculation. The dose of anti-VEGF antibody was 100 μ g twice weekly. Untreated animals received PBS alone. n, Necrotic areas. Scale bar, 100 μ m.



various molecular species of VEGF. But most of the endothelial cell mitogenic activity released by the tumour cells *in vitro* was strongly retained by heparin-sepharose and was eluted in the presence of 0.9 M NaCl (unpublished observations). This chromatographic behaviour is consistent with VEGF₁₆₅ but not with the other isoforms of VEGF¹⁶. This molecular species of VEGF is soluble after secretion and therefore is largely free to diffuse and reach its receptors in the vasculature^{6,16}.

Our findings demonstrate for the first time, to our knowledge, that blocking the action of a paracrine mediator that acts on the vasculature may have a significant or even dramatic inhibitory effect on tumour growth and emphasize the significance of VEGF as an important mediator of tumour angiogenesis. Therefore, blocking VEGF action has the potential to be of therapeutic significance for several highly vascularized and aggressive malignancies. □

Received 21 January; accepted 8 March 1993.

1. Klagsbrun, M. & D'Amore, P. A. *Rev. Physiol.* **53**, 217-239 (1991).
2. Folkman, J. & Shing, Y. *J. Biol. Chem.* **267**, 10931-10934 (1992).
3. Folkman, J. *J. Natn. Cancer Inst.* **82**, 4-6 (1991).
4. Ingber, D. *et al. Nature* **348**, 555-557 (1990).
5. Weidner, N., Semple, P., Welch, W. & Folkman, J. *N. Engl. J. Med.* **324**, 1-5 (1991).
6. Ferrara, N., Houck, K. A., Jakeman, L. B. & Leung, D. W. *Endocr. Rev.* **13**, 18-32 (1992).
7. Ferrara, N. & Henzel, W. J. *Biochem. biophys. Res. Commun.* **161**, 851-858 (1989).

8. Plouët, J., Schilling, J. & Gospodarowicz, D. *EMBO J.* **8**, 3801-3806 (1989).
9. Conn, G. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 2628-2632 (1990).
10. Jakeman, L. B., Winer, J., Bennett, G. L., Altar, C. A. & Ferrara, N. *J. clin. Invest.* **89**, 244-253 (1992).
11. Connolly, D. T. *et al. J. Biol. Chem.* **264**, 20017-20024 (1989).
12. Keck, P. J. *et al. Science* **246**, 1309-1311 (1989).
13. Leung, D., Cachianes, G., Kuang, W.-J., Goeddel, D. V. & Ferrara, N. *Science* **246**, 1306-1309 (1989).
14. Houck, K. A. *et al. Molec. Endocr.* **5**, 1806-1814 (1991).
15. Trisher, E. *et al. J. Biol. Chem.* **266**, 11947-11954 (1991).
16. Houck, K. A., Leung, D. W., Rowland, A. M., Winer, J. & Ferrara, N. *J. Biol. Chem.* **267**, 26031-26037 (1992).
17. Phillips, H. S., Hains, J., Leung, D. W. & Ferrara, N. *Endocrinology* **127**, 965-967 (1990).
18. Breier, G., Albrecht, U., Sterrer, S. & Risau, W. *Development* **114**, 521-528 (1992).
19. Senger, D., Perruzzi, C. A., Feder, J. & Dvorak, H. F. *Cancer Res.* **46**, 5629 (1986).
20. Rosenthal, R., Megyesi, J. F., Henzel, W. J., Ferrara, N. & Folkman, J. *Growth Factors* **4**, 53-59 (1990).
21. Berse, B., Brown, L. F., Van de Water, L., Dvorak, H. F. & Senger, D. R. *Molec. Biol. Cell* **3**, 211-218 (1992).
22. Shweiki, D., Itin, A., Soffer, D. & Keshet, E. *Nature* **359**, 843-845 (1992).
23. Plate, K. H., Breier, G., Weich, H. A. & Risau, W. *Nature* **359**, 845-847 (1992).
24. Berkman, R. A. *et al. J. clin. Invest.* **91**, 153-159 (1993).
25. Ferrara, N. *et al. J. clin. Invest.* **91**, 160-170 (1993).
26. Kim, K. J., Li, B., Houck, K., Winer, J. & Ferrara, N. *Growth Factors* **7**, 53-64 (1992).
27. Lieberman, T. A. *et al. EMBO J.* **6**, 1627-1632 (1987).
28. Schweigerer, L. *et al. Proc. natn. Acad. Sci. U.S.A.* **84**, 842-846 (1987).
29. Rizzino, A. *Meth. Enzym.* **146**, 341-348 (1987).
30. Westphal, M. *et al. Expl. Cell Biol.* **55**, 152-161 (1987).
31. Berman, P. W. *et al. J. Virol.* **63**, 3489-3498 (1989).
32. Ravindranath, N., Little-Ihrig, L., Phillips, H. S., Ferrara, N. & Zeleznick, A. J. *Endocrinology* **131**, 254-260 (1992).

ACKNOWLEDGEMENTS. We thank M. Westphal and W. Hamel for the generous gift of the G55 cell line, R. Shalaby for his help and advice with nude mice experiments, W. Young and J. Silva for excellent animal work, D. Finkle for histology, C. Kyle for immunocytochemistry and N. Singh for ELISA, D. Giltman and E. Glickerson for statistical analysis, J. Baker, G. Vehar, M. Cronin and R. Thomas for reading the manuscript and B. Fendly and E. Patzer for helpful discussions and advice.

Recombinant fibroblast growth factor-1 promotes intimal hyperplasia and angiogenesis in arteries *in vivo*

Elizabeth G. Nabel*, Zhi-yong Yang*, Gregory Plautz*, Reza Forough†, Xi Zhan†, Christian C. Haudenschild‡, Thomas Maciag† & Gary J. Nabel*

* Departments of Internal Medicine and Biological Chemistry, Howard Hughes Medical Institute, University of Michigan Medical Center, Ann Arbor, Michigan 48109-0650, USA
 † Departments of † Molecular Biology and ‡ Experimental Pathology, Holland Laboratory, American Red Cross, Rockville, Maryland 20855, USA

THE prototype members of the heparin-binding fibroblast growth factor (FGF) family¹⁻⁶, acidic FGF (FGF-1) and basic FGF (FGF-2), are among the growth factors that act directly on vascular cells to induce endothelial cell growth and angiogenesis. *In vivo*, the role of the FGF prototypes in vascular pathology has been difficult to determine. We report here the introduction, by direct gene transfer into porcine arteries, of a eukaryotic expression vector encoding a secreted form of FGF-1. This somatic transgenic model defines gene function in the arterial wall *in vivo*. FGF-1 expression induced intimal thickening in porcine arteries 21 days after gene transfer, in contrast to control arteries transduced with an *Escherichia coli* β -galactosidase gene. Where there was substantial intimal hyperplasia, neocapillary formation was detected in the expanded intima. These findings suggest that FGF-1 induces intimal hyperplasia in the arterial wall *in vivo* and, through its ability to stimulate angiogenesis in the neointima, FGF-1 could stimulate neovascularization of atherosclerotic plaques. Potentially, gene transfer of FGF-1 could also be used as a genetic intervention to improve blood flow to ischaemic tissues in selected clinical settings.

The FGF prototypes lack a classic signal sequence¹ (ss) for secretion, making it difficult to study their biological effects as extracellular polypeptides. It is now possible, however, to deliver recombinant genes directly into vascular cells at specific sites

*in vivo*⁷⁻¹² to determine their effects in the arterial wall. A secreted form of the FGF-1 gene was derived by ligation of the signal sequence from the hst/KS3 (FGF-4) gene to the 5' end of the open reading frame of FGF-1 (ref. 13) in the pMEX neo eukaryotic expression vector¹⁴. The pMEX neo-ss-hst/KS3:FGF-1 expression vector plasmid was transfected into porcine iliofemoral arteries by direct gene transfer^{8,15,16}, and controls were transfected with the *E. coli* β -galactosidase gene. The presence of the ss-hst/KS:FGF-1 plasmid was confirmed using polymerase chain reaction (PCR) in transfected iliofemoral arterial segments (Fig. 1a, lanes 1 and 2) but not in nontransfected carotid artery segments from the same pig (data not shown), and the presence of its messenger RNA was confirmed by reverse transcription PCR (Fig. 1b, lane 4). Expression of recombinant FGF-1 protein was confirmed by immunohistochemistry in transduced arterial segments. Porcine arteries transfected with ss-hst/KS:FGF-1 had immunoreactive protein primarily in the intima, including the endothelium, 21 days after transfection, whereas no FGF-1 protein was detected in arteries transduced with the β -galactosidase expression vector (Fig. 2a-c).

To evaluate the response of the arterial wall to expression of ss-hst/KS:FGF-1, the transfected artery segments were examined by light microscopy 21 days after gene transfer. Animals transduced with β -galactosidase showed minimal intimal thickening in iliofemoral artery segments, in contrast to the ss-hst/KS:FGF-1-transduced arteries (compare Fig. 3a and b). By quantitative morphometry, the intimal to medial ratio was more than sixfold greater in FGF-1 than β -galactosidase-transduced vessels (0.27 ± 0.06 versus 0.04 ± 0.01 , $P = 0.003$). Finally, in several experimental subjects, expression of ss-hst/KS:FGF-1 induced the formation of capillaries in the neointima (Fig. 4a, b), an effect not observed with the control.

Thus expression of secreted recombinant FGF-1 induced significant intimal proliferation and angiogenesis *in vivo*. Angiogenic factors^{1,2,6} have been classified previously into two categories: those that act directly on vascular endothelial cells to stimulate locomotion and mitosis and those that act indirectly to induce host cells to release growth factors that target the endothelial cell. In addition, because the FGF prototypes lack a classic signal sequence for secretion, their normal mode of release is not fully understood. They are detected after arterial

injury^{17,18} and can be found in the subendothelial matrix¹⁹. Indeed, NIH 3T3 cells can release FGF-1 *in vitro* in response to heat shock²⁰. In an injury model *in vivo*, systemically administered FGF-2 is a potent mitogen for vascular smooth muscle cells²¹, but it has not previously been possible to deter-

mine the function of the FGF prototypes in the normal arteries. Analysis of tissue sections in this study suggests that mitotic activity is induced by FGF-1 in the vessel wall, although it remains possible that FGF-1 affects cell migration, as suggested for platelet-derived growth factor (PDGF)²².

Several recombinant genes induce intimal smooth muscle hyperplasia, including PDGF¹⁶ or TGF- β 1 (E.G.N. *et al.*, unpublished results). In contrast, FGF-1 stimulates intimal hyperplasia and also induces the formation of new blood vessels in the expanded intima. These data therefore suggest that smooth muscle hyperplasia alone is not sufficient for the formation of new capillaries. Luminal endothelial cells in the iliofemoral

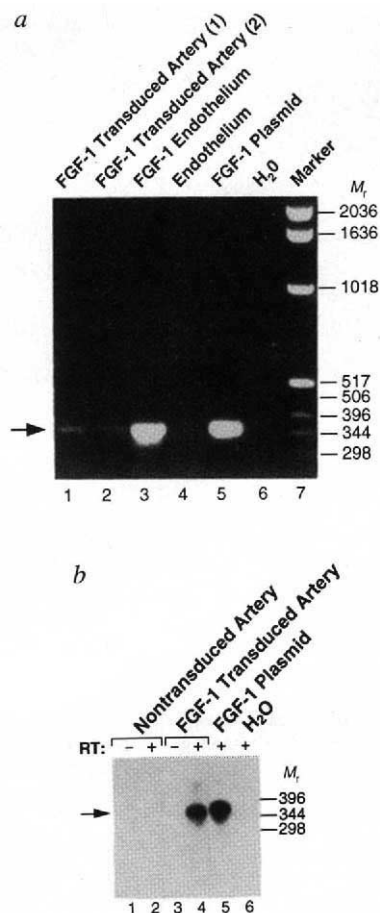


FIG. 1 Presence of ss-hst/KS:FGF-1 plasmid DNA (a) and mRNA (b) in porcine arteries after direct gene transfer *in vivo*. a, Recombinant ss-hst/KS:FGF-1 DNA was transfected into porcine right (lane 1) and left (lane 2) iliofemoral arteries and was detected by PCR 7 days after direct gene transfer using ethidium staining of agarose gels. *In vitro* transfected porcine endothelial cells (FGF-1 endothelium) (lane 3) compared with nontransduced porcine endothelial cells (Endothelium) (lane 4) were analysed as positive and negative controls. Additional controls included the recombinant FGF-1 plasmid (lane 5) and water (lane 6). b, The presence of ss-hst/KS:FGF-1 mRNA was detected using reverse transcription PCR by methods described previously¹⁶. The presence of recombinant FGF-1 mRNA was analysed by Southern blotting of PCR-amplified complementary DNA in nontransduced or transduced arteries treated with or without the addition of reverse transcriptase (RT) as indicated for lanes 1–4. Recombinant FGF-1 plasmid (lane 5) and water (lane 6) were included as positive and negative controls, respectively.

METHODS. Recombinant FGF-1 gene transfer in arterial segments was analysed by PCR of genomic DNA as previously described¹⁵. To conduct PCR analysis of recombinant FGF-1 transfected vessels, primers were synthesized from the cDNA sequence²⁰ which generated a 364 base pair (bp) fragment: sense (25 mer): CAA ACT CCT CTA CTG TAG CAA CCG G; antisense (25 mer): TTG CTT TCT GGC CAT AGT GAG TCC G. The sense primer was selected from a region 50 bp upstream to the transcription start site. Samples were analysed by ethidium bromide staining on a 1% agarose gel. The ss-hst/KS:FGF-1 chimera was prepared and inserted into the pMEX neo expression vector as previously described²⁰. Primary porcine endothelial muscle cell cultures were established as previously described and transfected with Lipofectin (BRL)¹⁵ to test the expression of the ss-hst/KS:FGF-1 vector. Transduced endothelial cells were assayed for secretion of recombinant FGF-1 into culture supernatants with a colorimetric proliferation assay by standard methods²⁴.

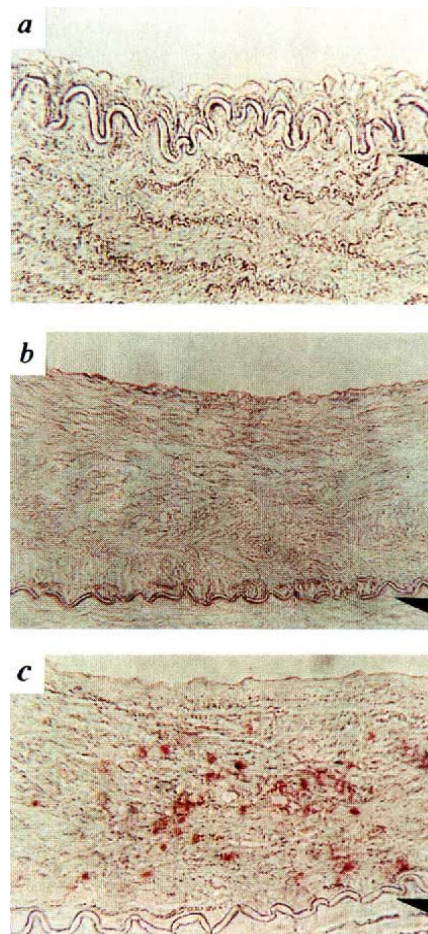


FIG. 2 Expression of recombinant ss-hst/KS:FGF-1 protein in porcine artery cells after direct gene transfer. Immunohistochemical staining of porcine arteries transduced with the *E. coli* β -galactosidase gene (a), or recombinant ss-hst/KS:FGF-1 gene (b, c) at 21 days using a control purified rabbit IgG antibody (b), or an affinity-purified rabbit antibody to FGF-1 (a, c). Arrow denotes the internal elastic lamina. (Magnification $\times 300$).

METHODS. Recombinant FGF-1 protein expression was analysed by immunohistochemistry of artery segments transduced with the recombinant ss-hst/KS:FGF-1 or *E. coli* β -galactosidase genes. Artery segments embedded in Tissue-Tek OCT compound (Miles, Elkhart, IN) were sectioned (6 μ m) and were incubated in 10 mM Tris-Cl pH 7.5, 150 mM NaCl, 1 mM EDTA (TNE), and 1% fetal bovine serum with 1:150 dilution of an affinity-purified rabbit anti-human FGF-1 antibody for 1 h at room temperature. Endogenous peroxidase activity was blocked by preincubation in TNE with 0.1% H₂O₂ for 45 min²⁵. Peroxidase-conjugated goat anti-rabbit IgG (H+L) antibody (Vector Laboratories, Burlingame, CA) (1:400) was added for 30 min at room temperature and samples were stained in 50 μ M sodium acetate (pH 5.0), 20 μ g ml⁻¹ 3-amino-9-ethyl-carbazole, and 0.015% H₂O₂ for 15 min. FGF-1-transduced arterial segments were also analysed with a control first antibody, purified rabbit IgG, and peroxidase conjugated goat anti-rabbit IgG (1:400) (Vector Laboratories).

FIG. 3 Intimal hyperplasia in porcine arteries 21 days after direct gene transfer of ss-hst/KSFGF-1 and *E. coli* β -galactosidase. Vessels were transduced with an *E. coli* β -galactosidase (negative control) (a) or ss-hst/KSFGF-1 (b) gene. Arrow denotes internal elastic lamina. (Magnification $\times 54$, haematoxylin-eosin stain).

METHODS. Direct intra-arterial gene transfer was done *in vivo* in 12 pigs, 6 with the recombinant ss-hst/KSFGF-1 gene and 6 with a control reporter gene, *E. coli* β -galactosidase. A double balloon intravascular catheter (C. R. Bard Inc., Billerica, MA) was inserted in porcine iliofemoral arteries as previously described⁸. The arterial segment isolated by the catheter was flushed with 5 ml saline and 5 ml Opti-MEM (BRL) to rinse blood from the vessel. The DNA liposome conjugates were prepared 10 min before transfection. Lipofectin (5 μ l) was diluted into 0.2 ml of Opti-MEM at room temperature, and 2–5 μ g plasmid DNA (stock concentration >1 mg ml⁻¹) was added. The solution remained at room temperature for 5–10 min, and 0.5 ml Opti-MEM was added. DNA liposomes were instilled into the arterial segment between the two balloons at 150 mm Hg in the left and right iliofemoral arteries and incubated for 20 min. Four of the twelve pigs were killed at one week, and eight pigs were killed at three weeks. Previous studies established that recombinant genes are stably expressed in vascular cells *in vivo* at 2–3 weeks^{8,15}, and intimal thickening induced by arterial manipulation is observed at this time point²⁶. Morphometric measurements of intimal and medial thickness were done in a blinded manner (by C.C.H.). Intimal-to-medial ratios were determined ($n=12$, $n=6$ *E. coli* β -galactosidase gene, $n=6$ FGF-1 gene) as previously described¹⁶. Intima-to-media ratios are expressed as a mean $\pm$ s.e.m. Comparisons between ss-hst/KSFGF-1 gene-transduced vessels and control vessels transduced with a reporter gene were made by two-tailed unpaired *t*-test.

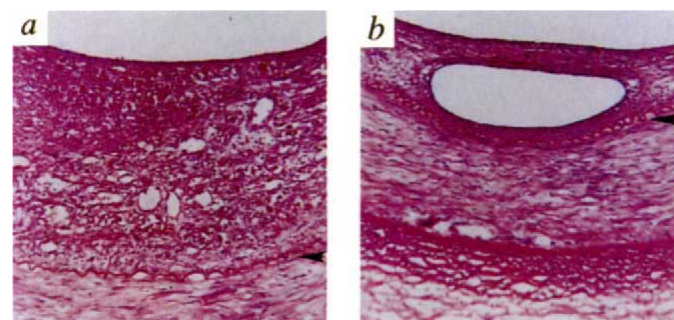
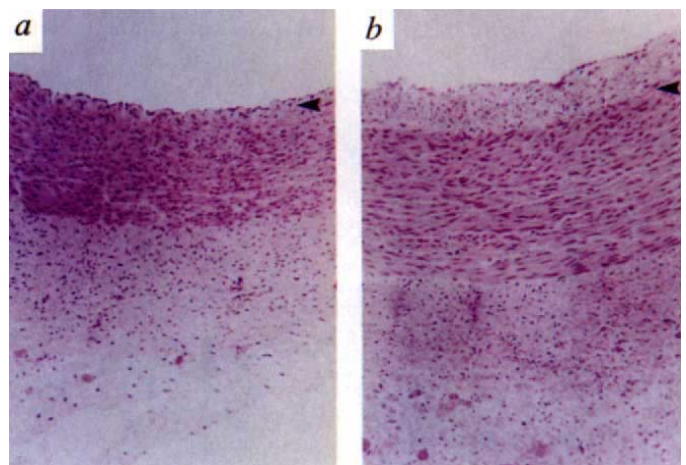


FIG. 4 Angiogenesis in the neointima of arterial segments after ss-hst/KSFGF-1 gene transfer. Vessels were transduced with ss-hst/KSFGF-1 and sections representing formation of multiple capillaries (a) or a larger intimal capillary (b) are shown. Arrow, Internal elastic lamina. (Magnification $\times 212$ (a), $\times 106$ (b), haematoxylin-eosin stain).

artery do not contain factor VIII, in contrast to those of the small capillaries in the adventitia. Endothelial cells lining the new intimal capillary beds are negative for von Willebrand factor (data not shown), suggesting that the FGF-1-induced capillaries arise from adjacent luminal endothelial cells. These findings are most consistent with a model in which FGF-1 acts locally on endothelial cells, perhaps through mitotic and locomotive effects. These angiogenic effects of FGF-1 could be mediated directly on endothelial cells and/or indirectly through the induction of other endogenous growth factors. In either case, the effects of FGF-1 are specific because they are not observed with

other recombinant growth factor genes^{15,16}. This model will aid the definition of such factors and the design of potential inhibitors of vascular cell proliferation.

The elaboration of FGF *in vivo* represents a potential mechanism to provide a blood supply to the hyperplastic intima. Such angiogenesis is observed in atherosclerotic plaques²³, and mechanisms to explain this observation have been lacking. Direct gene transfer of this FGF-1 or related genes *in vivo* could provide a method to stimulate collateral blood flow which would be beneficial for the treatment of ischaemic cardiovascular diseases. □

Received 21 October 1992; accepted 18 February 1993.

- Burgess, W. H. & Maciag, T. A. *Rev. Biochem.* **58**, 575–606 (1989).
- Ausprunk, D. H. & Folkman, J. *Microvasc. Res.* **14**, 53–65 (1977).
- Abraham, J. A. *et al. Science* **233**, 545–548 (1986).
- Gospodarowicz, D., Chen, J., Lui, G. M., Baird, A. & Bohlent, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 6963–6967 (1984).
- Thomas, K. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 6409–6413 (1985).
- Yanagisawa-Miwa, A. *et al. Science* **257**, 1401–1403 (1992).
- Nabel, E. G., Plautz, G., Boyce, F. M., Stanley, J. C. & Nabel, G. J. *Science* **244**, 1342–1344 (1989).
- Nabel, E. G., Plautz, G. & Nabel, G. J. *Science* **249**, 1285–1288 (1990).
- Plautz, G., Nabel, E. G. & Nabel, G. J. *Circulation* **83**, 578–583 (1991).
- Lim, C. S. *et al. Circulation* **83**, 2007–2011 (1991).
- Chapman, G. D. *et al. Circ. Res.* **71**, 27–33 (1992).
- Leclerc, G., Gal, D., Takeshita, S., Nikol, S., Weir, L. & Isner, J. M. *J. clin. Invest.* **90**, 936–944 (1992).
- Forough, R. *et al. J. biol. Chem.* **268**, 2960–2968 (1993).
- Martin-Zanca, D., Oskam, R., Mitra, G., Copeland, T. & Barbacid, M. *Molec. cell. Biol.* **9**, 24–33 (1989).
- Nabel, E. G., Plautz, G. & Nabel, G. J. *Proc. natn. Acad. Sci. U.S.A.* **89**, 5157–5161 (1992).

- Nabel, E. G. *et al. J. clin. Invest.* **91**, 1822–1829 (1993).
- Gajdusek, C. M. & Carbon, S. *J. Cell Physiol.* **139**, 570–579 (1989).
- McNeil, P. L., Muthukrishnan, L., Warder, E. & D'Amore, P. A. *J. Cell Biol.* **109**, 811–822 (1989).
- Vlodavsky, I., Folkman, J., Sullivan, R., Friedman, R. & Ishai-Michaeli, R. *Proc. natn. Acad. Sci. U.S.A.* **84**, 2292–2296 (1987).
- Jackson, A. *et al. Proc. natn. Acad. Sci. U.S.A.* **99**, 10691–10695 (1992).
- Lindner, V., Lappi, D. A., Baird, A., Majack, R. A. & Reidy, M. A. *Circ. Res.* **68**, 106–113 (1991).
- Jawien, A., Bowen-Pope, D. F., Lindner, V., Schwartz, S. M. & Clowes, A. *J. clin. Invest.* **89**, 507–511 (1992).
- Barger, A. C., Beeuwkes, R. III, Lainey, L. L. & Silverman, K. J. *N. Engl. J. Med.* **310**, 175–177 (1984).
- Mosmann, T. *J. immun. Meth.* **65**, 55–63 (1983).
- Ellis, S. A., Taylor, C. & McMichael, A. *Hum. Immun.* **5**, 49–59 (1982).
- Clowes, A. W., Clowes, M. M. & Reidy, M. A. *Lab Invest.* **54**, 295–303 (1986).

ACKNOWLEDGEMENTS. We thank D. Gordon for comments and advice, C. Enger for providing catheters and D. Gschwend for secretarial assistance. These studies were supported by grants from the NIH and the American Heart Association.

p53 is required for radiation-induced apoptosis in mouse thymocytes

Scott W. Lowe*, Earlene M. Schmitt*,
Sallie W. Smith†, Barbara A. Osborne†‡
& Tyler Jacks*§

* Department of Biology, Center for Cancer Research, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, USA and

† Department of Veterinary and Animal Sciences

‡ Program in Molecular and Cellular Biology, 309 Paige Laboratory, University of Massachusetts, Amherst, Massachusetts 01003, USA

§ To whom correspondence should be addressed

THE *p53* tumour suppressor gene is the most widely mutated gene in human tumorigenesis^{1,2}. *p53* encodes a transcriptional activator³⁻⁷ whose targets may include genes that regulate genomic stability^{8,9}, the cellular response to DNA damage^{10,11}, and cell-cycle progression^{12,13}. Introduction of wild-type *p53* into cell lines that have lost endogenous *p53* function can cause growth arrest¹⁴⁻¹⁶ or induce a process of cell death known as apoptosis^{17,18}. During normal development, self-reactive thymocytes undergo negative selection by apoptosis¹⁹, which can also be induced in immature thymocytes by other stimuli, including exposure to glucocorticoids¹⁵ and ionizing radiation¹⁶. Although normal negative selection involves signalling through the T-cell receptor¹⁴, the induction of apoptosis by other stimuli is poorly understood. We have investigated the requirement for *p53* during apoptosis in mouse thymocytes. We report here that immature thymocytes lacking *p53* die normally when exposed to compounds that may mimic T-cell receptor engagement and to glucocorticoids but are resistant to the lethal effects of ionizing radiation. These results demonstrate that *p53* is required for radiation-induced cell death in the thymus but is not necessary for all forms of apoptosis.

Because of the potential involvement of *p53* in inducing apoptosis^{17,18} we studied cell death in thymocytes derived from mice carrying a germ-line disruption of the *p53* gene (T.J. and R. Weinberg, unpublished results). Thymocytes were isolated from *p53* homozygous mutant, heterozygous and wild-type

animals and subjected *in vitro* to treatments that induce apoptosis. Although treatment with phorbol ester/calcium ionophore (which may mimic engagement of the T-cell receptor²²) and dexamethasone induced death with similar kinetics in thymocytes of all three genotypes (Fig. 1a, b), *p53*-deficient cells displayed a dramatic resistance to the effects of ionizing radiation (Fig. 1c). Moreover, the *p53*-deficient thymocytes remained viable following doses up to 2,000 centiGreys (cGy); wild-type cells were susceptible to treatment with as little as 100 cGy (Fig. 2). In addition, at all doses and times examined, cells isolated from heterozygous animals displayed intermediate viability compared to wild-type and homozygous mutant animals (Figs 1c and 2).

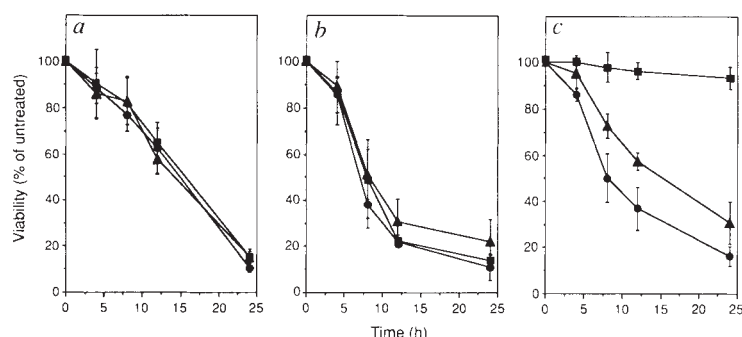
We further tested the effects of dexamethasone and γ -irradiation on thymocyte survival *in vivo*. Thymocytes isolated from treated animals were examined for the presence of the cell-surface markers CD4 and CD8 using two-colour fluorescence-activated cell sorter (FACS) analysis. Thymuses from untreated normal and mutant animals contained about 75–80% immature, CD4⁺CD8⁺ cells, which are susceptible to apoptosis²³. Forty-eight hours after treatment with dexamethasone, all thymuses sustained a significant reduction in cell numbers, which could be attributed to selective loss of CD4⁺CD8⁺ cells (Fig. 3a, b). Similarly, thymuses from wild-type animals exposed to γ -radiation contained a low percentage of CD4⁺CD8⁺ cells (Fig. 3a, b). In contrast, irradiation of *p53* homozygous mutant animals caused only minor reductions in CD4⁺CD8⁺ cells (Fig. 3a, b).

Conditions that induced cell death produced the internucleosomal degradation of thymic DNA, which is indicative of apoptosis²⁴ (Fig. 3c). This characteristic 'DNA ladder' was not evident after irradiation of homozygous mutant animals (Fig. 3c). Consistent with the data from *in vitro* experiments, mice that were heterozygous for the *p53* mutation were less susceptible than wild-type mice to the effects of ionizing radiation, both in survival of CD4⁺CD8⁺ cells and extent of DNA laddering (Fig. 3b, c).

Given the apparent requirement for *p53* function in radiation-induced apoptosis of thymocytes, we examined the steady-state level of *p53* protein in wild-type cells after exposure to ionizing radiation. Consistent with findings in other cell types^{10,11}, irradiation of thymocytes caused a dramatic increase in *p53* levels. The accumulation of *p53* protein was apparent within 1 h, before significant degradation of DNA (Fig. 4 and data not shown).

FIG. 1 Induction of apoptosis in isolated thymocytes. Thymocytes were treated with: a, 10nM phorbol ester (phorbol 12-myristate 13-acetate, PMA) and 500 nM calcium ionophore (A23187); b, 1 μ M dexamethasone; or c, 500 cGy ionizing radiation and viability was assessed at various times thereafter. Thymocytes were derived from *p53* homozygous mutant (■), heterozygous (▲) and wild-type (●) animals.

METHODS. Mutant mice used in these experiments carry a germ-line disruption of the *p53* gene (T.J. and R. Weinberg, unpublished results) that was made by gene targeting in D3 embryonic stem cells³². The mutation consists of a replacement by the bacterial *neo* gene of *p53* sequences between exons 2 and 6; immunoprecipitation analysis has confirmed that the mutation eliminates production of *p53* protein⁸. The mutation is carried on a hybrid (C57BL/6 $\times$ 129/sv) genetic background. Although genetic background influences the sensitivity of thymocytes to treatments such as irradiation³³, we obtained consistent results from all animals within a given genotype. Thymocytes were isolated from mice age 4.5–7 weeks in tissue culture media (DME supplemented with 5% fetal bovine serum and 25 mM HEPES, pH 7.2) and adjusted to a density of 1×10^6 per ml. At time zero, cultures were treated, divided into 16-mm wells (1×10^6 per well) and incubated at 37 °C. The relative amounts of nonviable cells were determined at various times by uptake of fluorescein isothiocyanate (FITC) and FACS analysis^{34,35}. Values represent the average



viability from four independent experiments with standard deviations; each experiment compared cells derived from one mutant, one heterozygote and one wild-type animal and were normalized to untreated samples from the same animal. Two experiments used littermates derived from F_1 crosses. There were no significant differences between viability of untreated thymocytes over the period of the experiment (not shown); on average, viability of untreated cells was about 70% at 24 h. Cell death by apoptosis was confirmed by analysis of genomic DNA (data not shown; see Fig. 3). Irradiation was done with a GammaCell 40 equipped with a ¹³⁷Cs source.

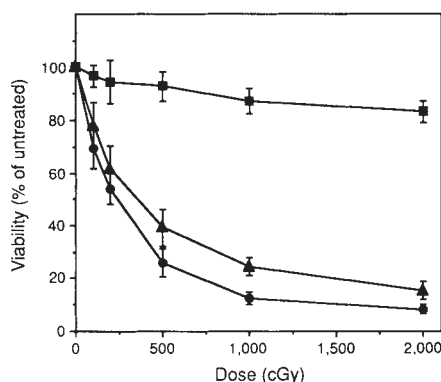


FIG. 2 Viability of isolated thymocytes treated with varying doses of ionizing radiation. Thymocytes were isolated from *p53* homozygous mutant (■), heterozygous (▲) and wild-type (●) animals and treated with varying doses of ionizing radiation. Viability was assessed 20 h after treatment as described in Fig. 1. Values represent averages from three independent experiments and are normalized to the amount of viable cells remaining in untreated cultures derived from the same animal.

In contrast, treatment with phorbol ester/calcium ionophore and dexamethasone resulted in little or no increase in *p53* levels.

These results establish the involvement of *p53* in a cell death pathway, specifically radiation-induced apoptosis in the thymus. Equally important, these data demonstrate that apoptosis can also occur in the absence of *p53* function. Thus, cell death in the thymus can be subdivided into at least two distinct pathways, one requiring *p53* and one that is *p53* independent. The existence of multiple apoptotic pathways in the thymus has been suggested from the analysis of *bcl-2* transgenic mice^{17,18}. Furthermore, the

apparently normal development of mice homozygous for a *p53* mutation (ref. 27 and T.J. and R. Weinberg, unpublished results) suggests that *p53* is not required for cell death in many, perhaps most, instances.

p53 has been implicated in controlling a checkpoint during the G1 phase of the cell cycle that may monitor the state of the DNA before entry into S phase^{11,28}. For example, *p53*-deficient fibroblasts fail to arrest transiently in G1 after γ -irradiation, although they still pause normally in G2 (ref. 11). A rapid accumulation of *p53* precedes G1 arrest in fibroblasts^{10,11} and, as shown here, radiation-induced apoptosis in thymocytes. Thus, the different cellular responses (apoptosis versus G1 arrest) may result from the activation of distinct target genes by *p53*. Alternatively, activation of the same target genes in the two cell types could have different consequences. The fact that elevated levels of *p53* can lead to the initiation of apoptosis is consistent with earlier studies that demonstrated a link between *p53* expression and cell death^{22,23}, and it is possible that many conditions that lead to an accumulation of *p53* could induce apoptosis. Those stimuli which cause apoptosis in thymocytes in the absence of *p53* function may use other transcription factors to activate the same set of 'cell death' genes.

The data presented here define another mechanism by which *p53* can act as a tumour suppressor gene. It has been proposed that the mutational inactivation of *p53* during tumorigenesis might allow the further accumulation of oncogenic mutations, due to the removal of an important G1 checkpoint^{11,28}. In thymocytes, and perhaps in other cell types as well, the absence

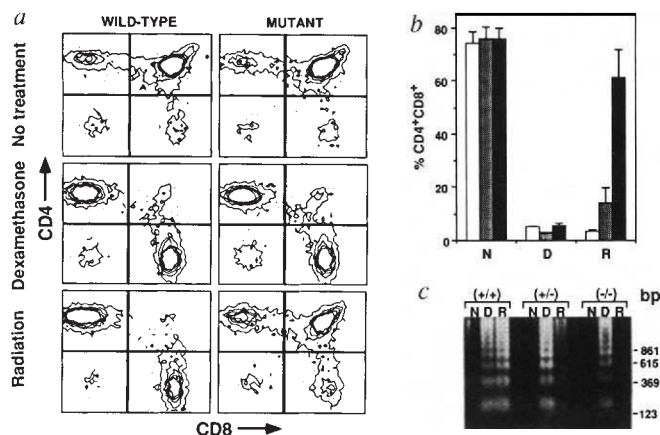


FIG. 3 Thymocyte apoptosis *in vivo*. *p53* homozygous mutant, heterozygous and wild-type animals were treated with dexamethasone or ionizing radiation and isolated thymocytes examined for the cell surface expression of CD4 and CD8 (after 48 h) and the condition of genomic DNA (after 10 h). *a*, Two-colour immunofluorescence contour plots from FACS analysis of CD4 and CD8 surface expression in wild-type and *p53* homozygous mutant mice. *b*, Mean percentage of surviving CD4⁺CD8⁺ thymocytes from wild-type (□), heterozygous (▨), and *p53* homozygous mutant (■) mice 48 h after treatment with dexamethasone (D), γ -irradiation (R), or no treatment (N). *c*, Agarose gel electrophoresis of total thymus DNA from wild-type (+/+), *p53* heterozygous (+/-), and *p53* homozygous mutant (-/-) mice 10 h after treatment with dexamethasone (D), γ -irradiation (R), or no treatment (N). The position of molecular size standards (in base pairs) is shown at right. **METHODS.** Thymocytes were recovered from mice 48 h after treatment with 0.5 mg dexamethasone (administered by intraperitoneal injection in PBS), γ -irradiation (500 cGy), or no treatment and stained with phycoerythrin-conjugated anti-CD4 and FITC-conjugated anti-CD8 antibodies (anti-L3T4 and anti-Lyt 2, Becton Dickinson). Multiparameter analysis of live cells was done on a FACStar Plus (Becton Dickinson). Dead cells were excluded by staining with propidium iodide and by gating of forward and side scatter of light during FACS analysis. The relative contributions of CD4 and CD8-expressing subpopulations were estimated using the Disp2D program (Becton Dickinson). Representative samples are shown in *a*. Values in *b* represent averages from three independent experiments (except the dexamethasone-treated heterozygous samples, which is an average of two experiments). Total genomic DNA was analysed from 10^6 thymocytes isolated 10 h after the treatments described above, according to the protocol in ref. 36.

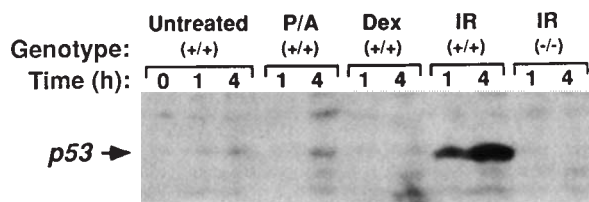


FIG. 4 *p53* levels in isolated thymocytes undergoing apoptosis. Thymocytes were isolated from wild-type (+/+) and homozygous mutant (-/-) animals and treated with cytotoxic agents. At 1 and 4 h after treatment, *p53* levels were determined by western blot analysis. The treatment and genotype of the cells are indicated above the appropriate lanes: normal tissue culture media (untreated); PMA + A23187 (P/A); dexamethasone (dex); ionizing radiation (IR). The time (in h) after treatment is indicated over each lane.

METHODS. Thymocytes were isolated and treated as described in Fig. 1 except the cultures were incubated in 75 cm² tissue culture flasks (10 ml per flask). For each sample, 1×10^7 cells were washed in PBS and lysed in Laemmli buffer³⁷. The proteins were separated on 7.5% SDS-polyacrylamide gels and transferred to PVDF membranes (Millipore). Membranes were blocked and probed with a pool of *p53*-specific monoclonal antibodies (PAb421, PAb240 and PAb248)^{38,39}. *p53* was detected using an alkaline phosphatase-conjugated secondary antibody and a chemiluminescent substrate (ref. 40; Lumi-Phos 530, Boehringer-Mannheim).

of *p53* function can lead to inappropriate cell survival after γ -irradiation. The failure to eliminate cells that have incurred DNA damage could lead to the selection of cells that have undergone neoplastic transformation. Note that among the various tumour types that occur in *p53* homozygous mutant mice, lymphoma is by far the most common (ref. 27 and T.J. and R. Weinberg, unpublished results), and the four cases of this tumour that have been examined from our *p53*-deficient mice have consisted predominantly of CD4⁺CD8⁺ cells (T.J., unpublished results). Thus, like *bcl-2* activation²⁹⁻³¹, the inactivation of *p53* may contribute to tumorigenesis through an inhibition of apoptosis. □

Received 30 December 1992; accepted 23 February 1993.

- Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. *Science* **253**, 49-53 (1991).
- Levine, A. J., Momand, J. & Finlay, C. A. *Nature* **352**, 453-456 (1991).
- Fields, S. & Jang, S. K. *Science* **249**, 1046-1049 (1990).
- Raycroft, L., Wu, J. & Lozano, G. *Science* **249**, 1049-1051 (1990).
- Farmer, G. et al. *Nature* **358**, 83-86 (1992).
- Kern, S. E. et al. *Science* **256**, 827-830 (1992).
- Zambetti, G. P., Bargonnetti, J., Walker, K., Prives, C. & Levine, A. J. *Genes Dev.* **6**, 1143-1152 (1992).
- Livingstone, L. R. et al. *Cell* **70**, 923-935 (1992).
- Yin, Y., Tainsky, M. A., Bischoff, F. Z., Strong, L. C. & Wahl, G. M. *Cell* **70**, 937-948 (1992).
- Kastan, M. B., Onyekwere, O., Sidransky, D., Vogelstein, B. & Craig, R. W. *Cancer Res.* **51**, 6304-6311 (1991).
- Kastan, M. et al. *Cell* **71**, 587-597 (1992).
- Michalovitz, D., Halevy, O. & Oren, M. *Cell* **62**, 671-680 (1990).
- Martinez, J., Georgoff, I., Martinez, J. & Levine, A. J. *Genes Dev.* **5**, 151-159 (1991).
- Mercer, W. E. et al. *Proc. natn. Acad. Sci. U.S.A.* **87**, 6166-6170 (1990).
- Diller, L. et al. *Molec. cell. Biol.* **10**, 5772-5781 (1990).
- Baker, S. J., Markowitz, S., Fearon, E. R., Willson, J. K. V. & Vogelstein, B. *Science* **249**, 912-915 (1990).
- Yonish-Rouach, E. et al. *Nature* **352**, 345-347 (1991).
- Shaw, P. et al. *Proc. natn. Acad. Sci. U.S.A.* **89**, 4495-4499 (1992).
- MacDonald, H. R. & Lees, R. K. *Nature* **343**, 642-644 (1990).
- Wyllie, A. H. *Nature* **343**, 555-559 (1990).
- Sellins, K. S. & Cohen, J. J. *J. Immun.* **139**, 3199-3206 (1987).
- Crabtree, G. R. *Science* **243**, 335-362 (1989).
- Smith, C. A., Williams, G. T., Kingston, R., Jenkinson, E. J. & Owen, J. J. T. *Nature* **337**, 181-184 (1989).
- Wyllie, A., Morris, R., Smith, A. & Dunlop, D. J. *Pathol.* **142**, 67-77 (1984).
- Sentman, C. L., Shutter, J. R., Hockenbery, D., Kanagawa, O. & Korsmeyer, S. *Cell* **67**, 879-888 (1991).
- Strasser, A., Harris, A. W. & Cory, S. *Cell* **67**, 889-899 (1991).
- Donehower, L. A. et al. *Nature* **356**, 215-221 (1992).
- Lane, D. P. *Nature* **358**, 15-16 (1992).
- McDonnell, T. J. et al. *Cell* **57**, 79-88 (1989).
- Strasser, A., Harris, A. W., Bath, M. L. & Cory, S. *Nature* **348**, 331-333 (1990).
- Tsujimoto, Y., Finger, L. R., Yunis, J., Nowell, P. C. & Croce, C. M. *Science* **226**, 1097-1099 (1984).
- Gossler, A., Doetschman, T., Korn, R., Serfling, E. & Kernier, R. *Proc. natn. Acad. Sci. U.S.A.* **83**, 9065-9069 (1986).
- Mori, N. et al. *Int. J. Radiat. Biol.* **62**, 153-159 (1991).
- Shi, Y. F. et al. *J. Immun.* **44**, 3326-3333 (1990).
- Butcher, E. & Ford, W. in *Handbook of Experimental Immunology* Vol. 2 57.13-57.17 (Blackwell Scientific, Oxford, 1986).
- Barry, M. A. & Eastman, A. *Arch. Biochem. Biophys.* **300**, 440-450 (1993).
- Harlow, E. & Lane, D. *Antibodies: a Laboratory Manual* (Cold Spring Harbor Laboratory Press, New York, 1988).
- Yewdell, J. W., Gannon, J. V. & Lane, D. P. *J. Virol.* **59**, 444-452 (1986).
- Gannon, J. V., Greaves, R., Iggo, R. & Lane, D. P. *EMBO J.* **9**, 1595-1602 (1990).
- Guarente, L. & Haldi, M. *J. biol. Chem.* **264**, 17107-17112 (1989).

ACKNOWLEDGEMENTS. We thank S. Conner and G. Paradis of the MIT Flow Cytometry Core Facility for expert assistance; D. Housman and R. Weinberg for support; L. Remington for technical assistance; D. Goldsby, M. Kastan, A. Strasser, L. M. Schwartz and P. Mombaerts for discussions and advice, and M. Weigel for help with the preparation of the manuscript. S.L. wishes to dedicate his contribution to this work to James Lowe. T.J. is a Lucille P. Markey Scholar. This work was supported by grants from the General Cinemas Corporation Charitable Foundation (T.J.) and the NIH (B.A.O.).

Thymocyte apoptosis induced by *p53*-dependent and independent pathways

A. R. Clarke, C. A. Purdie*, D. J. Harrison, R. G. Morris, C. C. Bird, M. L. Hooper & A. H. Wyllie

Cancer Research Campaign Laboratories, Department of Pathology, University Medical School, Teviot Place, Edinburgh EH8 9AG, UK

DEATH by apoptosis is characteristic of cells undergoing deletion during embryonic development, T- and B-cell maturation and endocrine-induced atrophy¹. Apoptosis can be initiated by various agents¹⁻⁵ and may be a result of expression of the oncosuppressor gene *p53* (refs 6-8). Here we study the dependence of apoptosis on *p53* expression in cells from the thymus cortex. Short-term thymocyte cultures were prepared from mice constitutively heterozygous or homozygous for a deletion in the *p53* gene introduced into the germ line after gene targeting. Wild-type thymocytes readily undergo apoptosis after treatment with ionizing radiation, the glucocorticoid methylprednisolone, or etoposide (an inhibitor of topoisomerase II), or after Ca²⁺-dependent activation by phorbol ester and a calcium ionophore. In contrast, homozygous null *p53* thymocytes are resistant to induction of apoptosis by radiation or etoposide, but retain normal sensitivity to glucocorticoid and calcium. The time-dependent apoptosis that occurs in untreated cultures is unaffected by *p53* status. Cells heterozygous for *p53* deletion are partially resistant to radiation and etoposide. Our results show that *p53* exerts a significant and dose-dependent effect in the initiation of apoptosis, but only when it is induced by agents that cause DNA-strand breakage.

We compared normal thymocytes with those from mice in which, as a result of gene targeting, a deletion was introduced to disable the *p53* gene. E14 embryonic stem (ES) cells, derived

from strain 129/Ola (ref. 9), and cultured in leukaemia inhibiting factor (LIF)-supplemented medium, were used to generate mutated clones carrying a deletion encompassing exons 2-6 of *p53*. Two targeting vectors were used, pCPR3.1 (Fig. 1a) and pCPR5.1. Both contained a 2.5-kilobase (kb) fragment from within intron 1, a 5-kb fragment including exons 7-11, and a *pgk-neo* cassette. In pCPR 5.1, a *pgk-tk* cassette was added at the 3' end of the genomic sequence. Following electroporation, 183 G418-resistant colonies and 204 colonies doubly resistant to G418 and Ganciclovir were screened by Southern blotting. Of these, a total of 5 clones carried the predicted modification to *p53*.

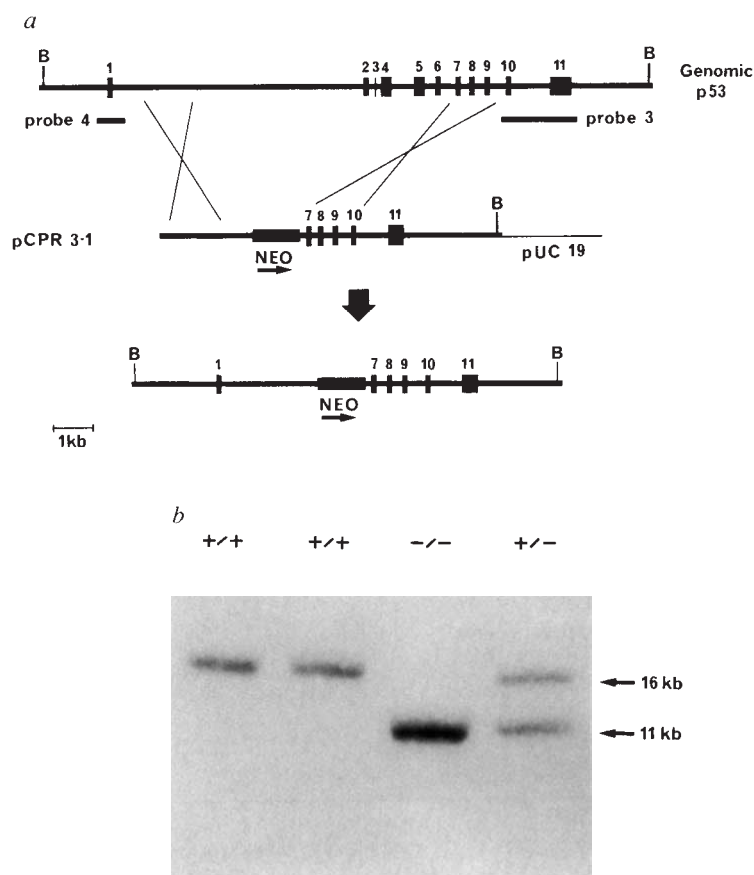
Germ-line chimaeras were obtained from one of these clones, R72, after injection into F₂ (C57BL/6×CBA) blastocysts. Offspring were analysed by Southern blotting of their DNA to confirm the presence of the mutated allele and mice heterozygous (+/-) for the mutation were mated to generate animals homozygous (-/-) at the mutant locus (Fig. 1b). Chimaeras were exclusively mated to strain 129/Ola females and therefore all experimental animals had an inbred 129/Ola genetic background.

Suspensions of thymocytes obtained from normal animals showed the expected sensitivity *in vitro* to γ -radiation, methylprednisolone, etoposide and a calcium-dependent activation stimulus provided by a combination of phorbol myristoyl 13-acetate (PMA) with the calcium ionophore ionomycin. Cells from *p53*-null animals, however, retained their sensitivity to glucocorticoid but were completely resistant to induction of apoptosis by radiation, even at the relatively high dose of 7 Gy (Fig. 2). This effect of *p53* was gene-dose-dependent, as shown by the intermediate sensitivity of cells from animals bearing only a single effective copy of *p53*. Thymocytes from homozygous null *p53* mice, and (to a slightly lesser extent) from heterozygous animals, were also relatively resistant to apoptosis induced by the topoisomerase II inhibitor, etoposide (Fig. 3a). Unlike radiation, however, etoposide at increasing concentrations elicited progressively more apoptotic cells. PMA with ionomycin was similar to glucocorticoid in being equally effective at inducing apoptosis in normal, heterozygous and

* Present address: Department of Pathology, University of Glasgow, Western Infirmary, Glasgow G11 6NT, UK.

FIG. 1 Strategy for the inactivation of the murine *p53* gene. *a*, The *p53* locus (boxes represent exons 1–11). The vector pCPR3.1 used to generate clone R72 was created by ligation of a *HindIII*–*Bam*HI fragment of intron 1 to a *Bam*HI–*Eco*RI fragment encompassing exons 7–11 (from murine genomic *p53* clone pMSVp53G; ref. 22), and then inserting a *pgk-neo* cassette²³ at the junction of these fragments. pCPR5.1 differs by the inclusion of a *pgk-tk* cassette at the *Bgl*III site at the 3' end of the *p53* homology (1.8 kb *Bgl*III–*Eco*RI HSV-*tk* fragment²⁴ in the *Sma*I site of a *pgk* expression vector²⁵). The two probes used for Southern analysis are indicated, one internal (P3, a 2-kb *Hind*III fragment) and one external (P4, a 0.8 kb fragment from a *Eco*RI–*Hind*III double digest) to the homology. Thick lines represent genomic DNA; thin lines, plasmid DNA. 'B' indicates a *Bgl*III restriction enzyme digestion site. *b*, Typical Southern analysis of progeny mice after restriction with *Bgl*III and probing with P3. An identical pattern is obtained with P4. The positions of bands corresponding to the wild-type *p53* allele (16 kb) and the *neo*-containing allele (11 kb) are indicated.

METHODS. Electroporation was done as described²⁶. After one day, targeted cells were selected in $200 \mu\text{g ml}^{-1}$ G418 and, when counterselection was required, in $2 \mu\text{M}$ Ganciclovir in the continued presence of G418 after 4 days. Clones were transferred into G418-containing medium after 10 days. Following identification of targeted clones and production and breeding of germ-line chimaeras, DNA was prepared from tail biopsies, digested, separated by electrophoresis in a 0.8% agarose gel, transferred to a Hybond-N⁺ filter (Amersham) and hybridized according to the manufacturer's instructions.



homozygous-null thymocytes (Fig. 3a); the rates at which these cells entered apoptosis in the absence of an inducing stimulus were comparable (Fig. 3b).

Expression of *p53* was also investigated immunocytochemically in smears of thymocytes from wild-type mice using the murine-specific antibodies CM-5 (a polyclonal antibody) and PAb242 (a monoclonal antibody against N-terminal epitopes of *p53*; ref. 10). Over the first four hours of culture, the proportion of etoposide- and radiation-treated cells showing intranuclear *p53* rose to 6.5%, whereas less than 3% of nuclei from untreated and glucocorticoid-treated thymocytes were positive.

These data raise several interesting issues regarding the role of the *p53* oncosuppressor gene in apoptosis. Previous studies have shown that expression of wild-type *p53* (but not of the mutated, oncogenic form) induces apoptosis in cells of erythroid⁶, myeloid⁷ and colorectal epithelial lineages⁸. However, in these experiments *p53* was expressed in a non-physiological environment as part of a recombinant construct and introduced into immortalized cells in which other uncharacterized changes may have occurred. Here, we have studied inactivation of one or both copies of the normal gene in otherwise normal cells from a fully differentiated organ. We show that *p53* has a limiting, gene-dose-dependent effect on apoptosis following some, but not all, types of lethal stimuli. By contrast, using thymocytes from animals heterozygous for an *Rb-1* null allele¹¹, we found no evidence for a similar role for the retinoblastoma oncosuppressor gene (data not shown).

Ionizing radiation and etoposide both differ from glucocorticoid in causing immediate DNA-strand breaks in thymocytes¹², which in the case of ionizing radiation are mainly single-strand nicks that are repaired within an hour¹³. The irradiated cells are nonetheless committed to apoptosis, which is initiated in a stochastic fashion over the ensuing several hours, suggesting that a lethal signal is generated within the damaged cells that is not revoked by subsequent DNA repair. Intranuclear *p53*

protein is known to accumulate following DNA damage^{14,15}, and our data provide evidence that this is part of the lethal signal.

Etoposide stabilizes the topoisomerase–DNA complex during its cleavage–religation cycle, and so generates double-strand DNA breaks, even in non-replicating cells¹⁶. The reason for the efficacy of such drugs in cancer chemotherapy is not clear, although they can initiate apoptosis in a variety of cell types^{12,17,18}. Our data show that the apoptosis caused by etoposide in cortical thymocytes, which are rich in topoisomerase II, depends largely upon *p53*, although additional modes of drug action may apply at higher concentrations.

There must also be a *p53*-independent pathway responsible for the apoptosis initiated by glucocorticoid treatment, calcium-dependent activation and *in vitro* ageing. The combination of calcium ionophore and the protein kinase C activator, PMA, reproduces many features of the 'activation-induced' death of rodent thymocytes achieved by crosslinking the T-cell antigen receptor¹⁹. Thus it provides a model of the thymocyte death involved in the physiological process of negative selection. Calcium-dependent activation and glucocorticoid sometimes exert mutually antagonistic effects on thymocyte apoptosis¹⁹, suggesting that this *p53*-independent pathway may itself be complex. It may involve protein synthesis, as thymocyte apoptosis induced by either ionophore or glucocorticoid is abrogated by the protein synthesis inhibitor cycloheximide²⁰. Cycloheximide also blocks the thymocyte apoptosis caused by radiation⁴ and topoisomerase II inhibitors¹², although it does not affect the immediate DNA breakage caused by these agents. The *p53*-dependent and independent pathways may therefore share common elements that require protein synthesis downstream of *p53*. The relationship of the pathways to each other might be resolved through study of *p53* substrates (such as the growth-arrest gene *GADD45*; ref. 21) that might be common to both.

In conclusion, we have demonstrated a strictly *p53*-dependent pathway to apoptosis in thymocytes, cells that are oriented

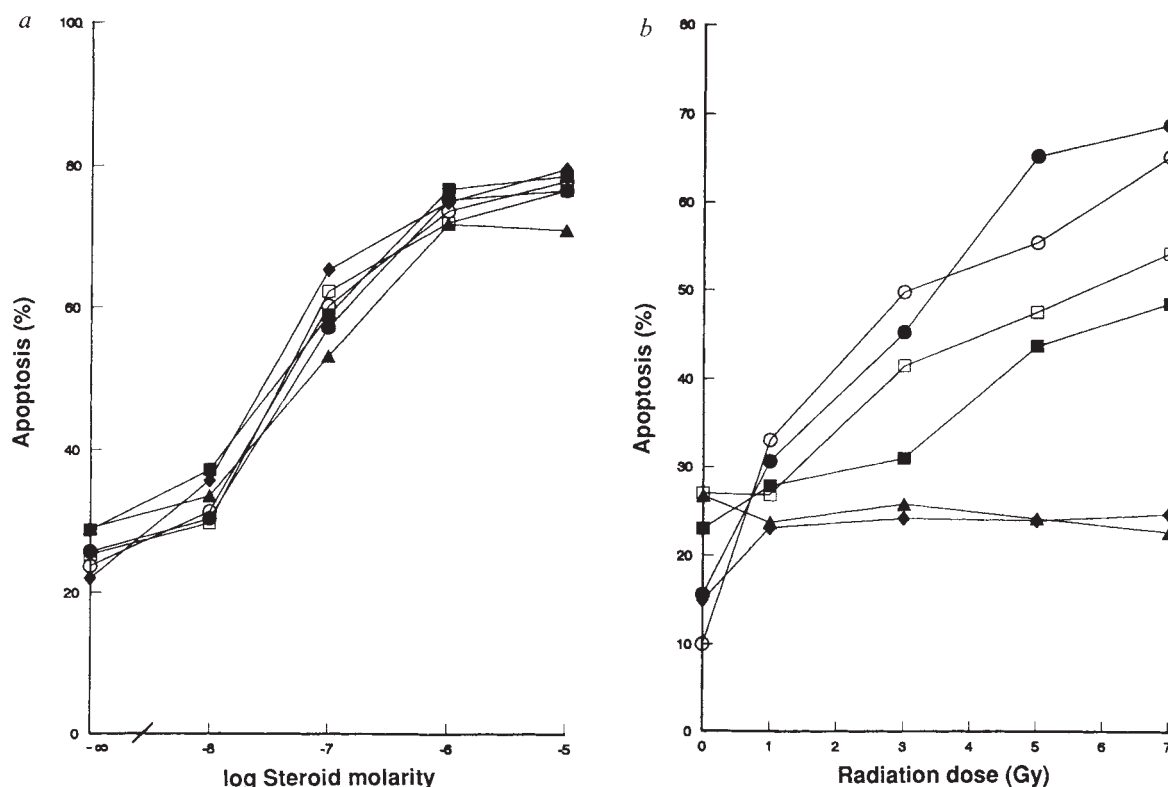


FIG. 2 Induction of apoptosis in thymocytes in suspension culture after 8 h exposure to methylprednisolone (a) or 8 h after γ -irradiation (b). Cells from the same thymus glands appear in both a and b, and each thymus is given a unique symbol according to its p53 genotype: wild-type (+/+), filled and open circles; heterozygote (+/-), filled and open squares; homozygous null (-/-), triangles and diamonds.

METHODS. Suspensions of thymocytes ($100 \mu\text{l}$) at 10^7 ml^{-1} in Glasgow-modified Eagle's medium (GMEM)³ were placed in the wells of a microtitre plate with methylprednisolone sodium succinate (Solu-medrone, Upjohn) diluted in isotonic saline, or with saline as control. For irradiation, immediately

before plating thymocytes were suspended in 1.5-ml screw-capped nylon tubes and exposed to γ -rays from a ^{137}Cs source at $\sim 0.5 \text{ Gy min}^{-1}$. After incubation of the plates at 37°C in 5% CO_2 in air, cells were fixed in 4% formaldehyde in ethanol overnight, washed three times in water, stained with $10 \mu\text{g ml}^{-1}$ acridine orange in saline and viewed for nuclear chromatin morphology in a fluorescence microscope. Apoptotic cells were unequivocally recognized by their condensed, sometimes fragmented chromatin¹. All data points are the means of triplicate wells. At least 200 cells were counted from each well, giving reproducibility to within 5%.

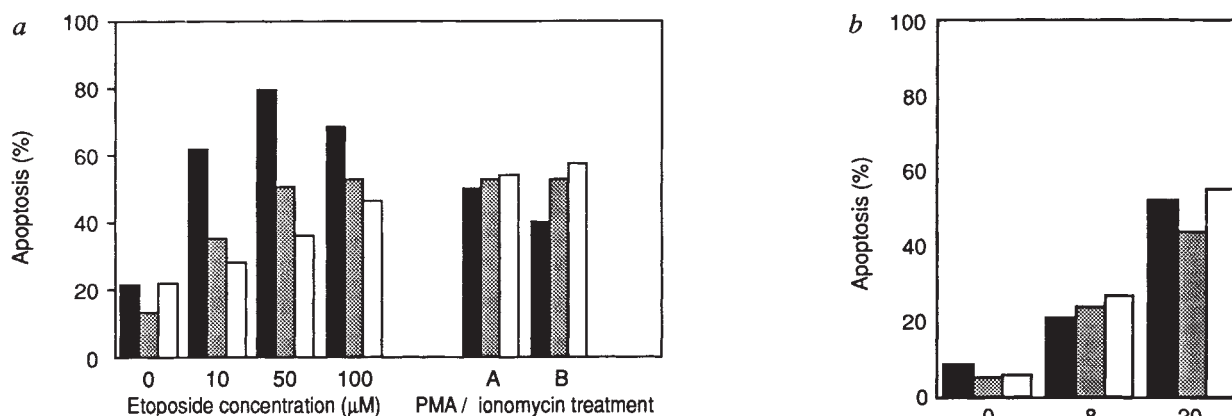


FIG. 3 Induction of apoptosis in thymocytes after 8 h exposure to etoposide or PMA and ionomycin (a), or with increasing time in untreated cultures (b). The p53 status of these thymocytes was +/+, black bars; +/-, grey bars; -/-, open bars. A single thymus of each type was the source of all the cells in each of a and b.

METHODS. As for Fig. 2. Etoposide (Sigma) was initially made up in dimethyl-

sulphoxide at 100 mM and diluted in saline. PMA and ionomycin (Sigma) were applied at respective concentrations of 5 ng ml^{-1} and $5 \mu\text{g ml}^{-1}$ (condition A) or at fivefold greater concentrations of both agents (condition B). The no-treatment control contained 0.1% (v/v) dimethylsulphoxide in saline.

towards programmed deletion by this mode of death. Agents that initiate DNA-strand breakage kill thymocytes by this pathway, but other lethal stimuli are effective in the absence of *p53*. The *p53*-independent stimuli include several that mimic physiological cell-deletion signals, namely glucocorticoid, calcium-associated activation and ageing *in vitro*. □

Received 1 March; accepted 24 March 1993.

- Arends, M. J. & Wyllie, A. H. *Int. Rev. exp. Path.* **32**, 223–254 (1991).
- Smith, C. A., Williams, G. T., Kingston, R., Jenkinson, E. J. & Owen, J. J. T. *Nature* **337**, 181–184 (1989).
- Wyllie, A. H. *Nature* **284**, 555–556 (1980).
- Sellins, K. S. & Cohen, J. J. *J. Immun.* **139**, 3199–3206 (1987).
- McConkey, D. J., Orrenius, S. & Jondal, M. *Immun. Today* **11**, 120–121 (1990).
- Ryan, J. J., Danish, R., Gottlieb, C. A. & Clarke, M. F. *Molec. cell. Biol.* **13**, 711–719 (1993).
- Yonish-Rouach, E. *et al. Nature* **352**, 345–347 (1991).
- Shaw, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **89**, 4495–4499 (1992).
- Hooper, M. L., Hardy, K., Handyside, A., Hunter, S. & Monk, M. *Nature* **326**, 292–295 (1987).
- Yewdell, J. W., Gannon, J. V. & Lane, D. P. *J. Virol.* **59**, 444–452 (1986).
- Clarke, A. R. *et al. Nature* **359**, 328–330 (1992).

- Walker, P. R. *et al. Cancer Res.* **51**, 1078–1085 (1991).
- Zhivotovsky, B. D., Seiliev, A. A. & Hanson, K. P. *Int. J. Rad. Biol.* **42**, 199–204 (1982).
- Kastan, M. B., Onyekwere, O., Sidransky, D., Vogelstein, B. & Craig, R. W. *Cancer Res.* **51**, 5304–5311 (1991).
- Kuerbitz, S. J., Plunkett, B. S., Walsh, M. V. & Kastan, M. B. *Proc. natn. Acad. Sci. U.S.A.* **89**, 7491–7495 (1992).
- Roy, C. *et al. Expl. Cell Res.* **200**, 416–424 (1992).
- Bertrand, R., Sarang, M., Jenkin, J., Kerrigan, D. & Pommier, Y. *Cancer Res.* **51**, 6280–6285 (1991).
- Fanidi, A., Harrington, E. A. & Evan, G. I. *Nature* **359**, 554–556 (1992).
- Iseki, R., Mukai, M. & Iwata, M. *J. Immun.* **147**, 4286–4292 (1991).
- Wyllie, A. H., Morris, R. G., Smith, A. L. & Dunlop, D. J. *J. Path.* **142**, 67–78 (1984).
- Kastan, M. B. *et al. Cell* **71**, 587–597 (1992).
- Bienz, B., Zakut-Houri, R., Givol, D. & Oren, M. *EMBO J.* **3**, 2179–2183 (1984).
- McBurney, M. W. *et al. Nucleic Acids Res.* **19**, 5755–5761 (1991).
- Colbere-Garapin, F., Chousterman, S., Horodniceanu, F., Kourilsky, P. & Garapin, A.-X. *Proc. natn. Acad. Sci. U.S.A.* **76**, 3755–3759 (1979).
- van der Lugt, N., Robanus Maandag, E., te Riele, H., Laird, P. W. & Berns, A. *Gene* **105**, 263–267 (1991).
- Kaster, K. R., Burgett, S. G., Rao, R. N. & Ignolia, T. D. *Nucleic Acids Res.* **11**, 6895–6911 (1983).

ACKNOWLEDGEMENTS. We thank A. Peter and L. Dobbie for technical assistance, J. Verth and his staff for animal care, D. Lane for antibody CM5, and M. Oren for pMSVp53G. This work was supported by programme grants from the Cancer Research Campaign to C.C.B., A.H.W. and M.L.H., and by a Medical Faculty Fellowship to C.A.P.

Solution structure of the POU-specific DNA-binding domain of Oct-1

Niek Dekker^{*†}, Michel Cox^{*}, Rolf Boelens^{*}, C. Peter Verrijzer[†], Peter C. van der Vliet^{†‡} & Robert Kaptein^{*‡}

^{*} Bijvoet Center for Biomolecular Research, Utrecht University, Padualaan 8, 3584 CH Utrecht, The Netherlands

[†] Laboratory for Physiological Chemistry, Utrecht University, Vondellaan 24^a, 3521 GG Utrecht, The Netherlands

THE transcription factor Oct-1 belongs to a family containing a POU DNA-binding domain. This bipartite domain is composed of a POU-specific domain (POU_s) and a POU-homeodomain (POU_h) connected by a flexible linker. The left half of the optimal POU binding site, the octamer ATGCAAT, is recognized by POU_s and the right half by POU_h. We have determined the solution structure of POU_s by nuclear magnetic resonance. It consists of four α -helices connected by short loops. Helices I and IV are in a parallel coiled-coil arrangement. The folding topology appears to be similar to that of the bacteriophage λ -repressor and 434 repressor. For the well defined parts of the protein (residues 1–71), the average root-mean square deviation for the backbone atoms is 0.9 Å. Based on the observed selective exchange broadening in the (¹⁵N, ¹H)-HMQC (heteronuclear multiple quantum coherence) spectrum of the POU_s-DNA complex we conclude that DNA-binding is mediated by helix III. We propose a model for the POU-DNA complex in which both recognition helices from the two subdomains have adjacent positions in the major groove.

POU transcription factors are involved in transcriptional regulation, development and cell differentiation^{1–4}. One member of this family, Oct-1, is a ubiquitously expressed protein involved in the regulation of housekeeping genes⁵, and adenovirus DNA replication.^{6,7} The functional organization of the protein is shown in Fig. 1a. The DNA-binding domain consists of two subdomains, which are autonomously folded structures and both display sequence-specific DNA-binding^{8,9}. We cloned and expressed an 80 amino-acid fragment containing POU_s, which is functionally active in sequence-specific DNA binding⁹. The amino-acid sequence is depicted in Fig. 1b.

The resonance assignment was accomplished using two-dimensional and heteronuclear three-dimensional nuclear magnetic resonance (NMR) spectra of unlabelled and ¹⁵N-labelled

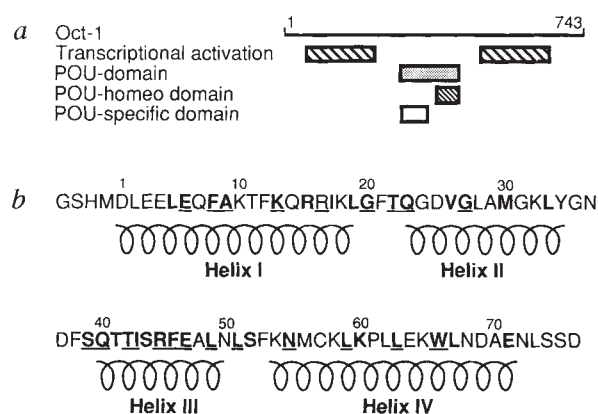


FIG. 1 a, Functional organization of the Oct-1 protein^{17,18}. b, Amino-acid sequence of the POU-specific DNA-binding fragment of Oct-1 used in this work. Residues both underlined and in bold, are absolutely conserved among the POU proteins, whereas residues only in bold, are subject to conservative changes (L/M, I/V/L, R/K, E/D, S/T and S/C; single-letter amino-acid code). The residues 1–71 entail the conserved POU_s domain (corresponding to residues 284 to 355 of Oct-1¹⁸), whereas residues 72–76 are part of the linker¹⁸. Residues –4 to –1 arise from the cloning procedure. The secondary structure elements (α -helices) are indicated below the sequence. METHODS. The protein was expressed by an *Escherichia coli* PET-15^b plasmid (Novagen). Purification was accomplished by Ni-NTA (Qiagen) affinity chromatography²⁰. After thrombin cleavage of the His-tagged protein, POU_s was purified by ion-exchange chromatography to homogeneity. Uniformly ¹⁵N-labelled POU_s was produced by using ¹⁵NH₄Cl as the sole nitrogen source in M9 minimal medium. NMR measurements were done on Bruker AMX 500 and AMX 600 spectrometers using 2–4 mM solutions of the protein in 0.1 M NaCl, 5 mM dithiothreitol, pH 5.0 at 25 °C.

POU_s (M.C. *et al.*, manuscript submitted). On the basis of the observed sequential and medium range connectivity patterns and the exchange data, four α -helices could be identified (Fig. 1b). From the nuclear Overhauser effect (NOE) spectra 53 intrasidual, 372 sequential, 328 medium-range and 263 long-range NOEs were obtained. Furthermore, 82 constraints for 41 hydrogen bonds, and 74 ϕ and 19 χ torsion angle constraints were determined. An overview of all interresidual NOE connectivities is given in Fig. 2.

The structure was determined on the basis of NMR data with the distance geometry program DGII (ref. 10). We used 1,098 upper distance bounds and 93 torsion angle constraints to calculate a set of 30 structures for POU_s, of which 23 had low final energy values after energy minimization and displayed no large violations (see Fig. 3a). The structures display an average root-mean square deviation (r.m.s.d.) about the mean coordinate

[‡] To whom correspondence should be addressed.

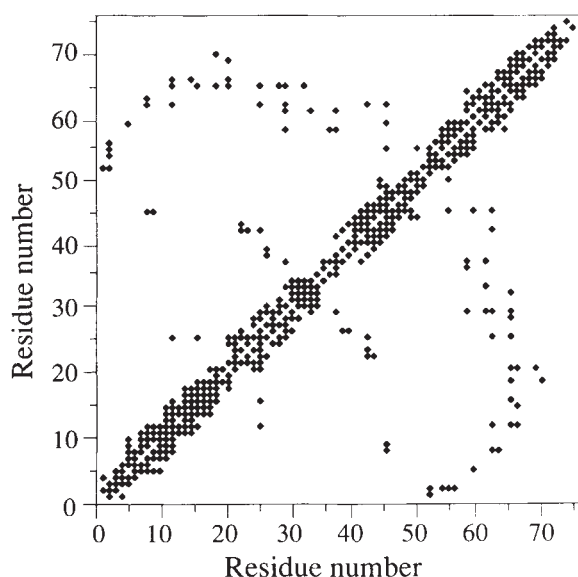


FIG. 2 Diagonal plot of inter-residual NOE constraints.

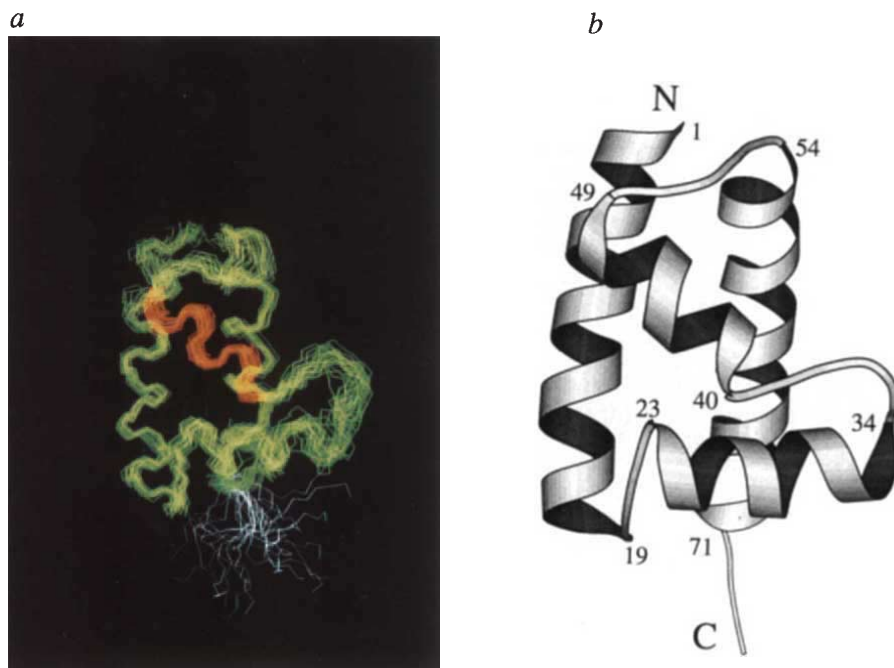
position of 0.9 Å for the backbone atoms and 1.6 Å for all heavy atoms for residues 1–71. In the well defined α -helical regions (comprising residues 7–29, 38–46 and 59–69) the r.m.s.d. is 0.6 Å for the backbone atoms.

The structure has a compact fold of four helices (Fig. 3b). Helices I (1–19) and IV (54–71) have a parallel orientation and are packed in a left-handed supercoil, similar to the coiled-coil arrangement seen in the crystal structure of the GCN4 leucine

zipper¹¹. Helix IV contains a proline residue at position 61, which disrupts the regular hydrogen-bond pattern, resulting in a kink of $\sim 15^\circ$. Helix II (23–34) is structurally less well defined, probably because of the presence of three glycines. Helix III (40–49) is involved in contacting the DNA (discussed below). Side chains of conserved amino acids of the four amphipathic helices make up the hydrophobic core of the protein. Aromatic residues that are absolutely conserved among the POU proteins (Phe 8, 46 and Trp 66) have well defined positions in the core; other conserved hydrophobic amino acids (Leu 5, 19, 51, 59 and 63, Val 26 and Ile 43) have buried positions in the structure. The absolutely conserved hydrophobic residues Ile 17 and Leu 49 are solvent exposed. It was found that the folding topology of POU₅ has a striking similarity with the first four helices of the bacteriophage repressors, λ -repressor and 434-repressor¹². Taking into account deletions at positions 32–37 and 53–54 of POU₅, the C α atoms of the first 60 residues of λ -repressor superimpose on POU₅ with an r.m.s.d. of 2.4 Å.

It is not known which residues of POU₅ are involved in DNA binding. Neither amino-acid conservation nor mutagenesis give any clue. To probe the DNA-binding site we developed a novel NMR approach. It is based on the idea that NH and NH₂ groups, that are involved in protein–DNA interactions, are likely to undergo relatively large chemical shift changes on complex formation. Alternatively, these protons may undergo enhanced solvent exchange because of the presence of the DNA. In both cases this may lead to selective signal loss, either due to line broadening or intensity reduction, that can be conveniently monitored in (¹⁵N, ¹H)-HMQC spectra of ¹⁵N enriched protein in the presence of small amounts of DNA. Indeed, as shown in Fig. 4a, selective signal loss was observed in the HMQC spectrum of the complex of ¹⁵N-labelled POU₅ with low amounts of the optimal POU₅ binding site (GAATATGCAA) present. This occurs for the amide side-chain resonances of Gln 23 and 40, of Asn 50 and 55 and for the N_H side chain resonances of Arg 45. At high salt conditions (0.7 M NaCl) the resonances

FIG. 3 *a*, The backbone traces (N, C α , C') of the 23 distance geometry conformers of POU₅ superimposed for residues 1–71; the recognition helix (III) is shown in red, and the flexible C-terminal tail (residues 72–76) is coloured white. *b*, Ribbon diagram representing the folding of POU₅ using the program Molscript²⁰. Numbers refer to the first and last residues in the helices. NOE connectivities from strong, medium and weak crosspeaks in the 600 MHz NOESY spectra with a 75 ms mixing time were assumed to correspond to proton–proton distances ≤ 2.8 , ≤ 3.5 and ≤ 4.5 Å apart, respectively. Additional NOE crosspeaks from NOE spectra with a 150 ms mixing time were assumed to be due to protons ≤ 5.5 Å apart. The NOE distances involving a methyl group were multiplied by a factor of 1.2 to correct for the three proton intensity. Standard values of pseudo-atom corrections²¹ were added except for methyl groups, for which a value of 0.3 Å was used²². All lower bounds were set to 2.0 Å. Pseudo atom corrections were not applied to residues Leu 2 and 59 and Phe 8 and 46, because these residues displayed a well defined orientation in preliminary structure calculations. Hydrogen bonds were constrained for 41 pairs between the amide proton and the carbonyl oxygen of the $i+4^{\text{th}}$ residue, by using constraints of $2.8 \leq d_{\text{NC}} \leq 3.2$ Å and $4.0 < d_{\text{NC}} < 4.3$ Å. Coupling constants $^3J_{\text{HN}\alpha}$ were estimated from splittings in the J -resolved HMQC spectra of the ¹⁵N-labelled protein²³, from which 74 ϕ angles could be classified to be $-65^\circ \pm 10^\circ$ or $-120^\circ \pm 30^\circ$. Stereospecific assignments²⁴ were obtained for the prochiral methyl groups of Val 23, which is in the gauche⁺ conformation. The χ_1 torsion angle constraints (19 in total) were obtained from the analysis of the $\alpha\beta$ coupling constants in the COSY spectrum and intraresidual NOE crosspeaks. Distance geometry calculations were done on Silicon Graphics



Personal Iris 4D/35 and 4D/120 GTX workstations with the DGLI program¹⁰ implemented in the Biosym NMRchitect software. Coordinates will be deposited in the Brookhaven Protein Data Bank. In the average structure 39 NOE constraints were violated with more than 0.4 Å, of which 11 with more than 0.7 Å (largest was 1.2 Å). Additionally, 3 dihedral constraints were violated with more than 10° (largest was 29°).

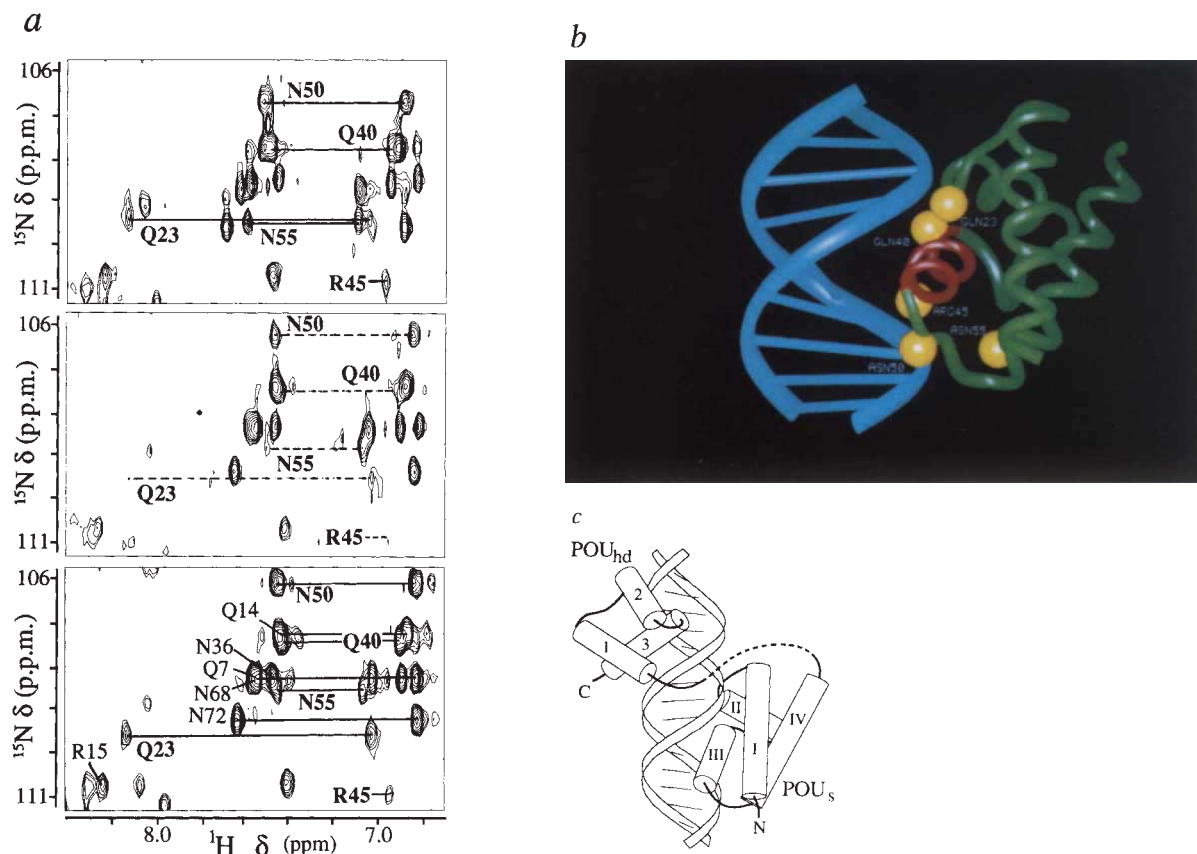


FIG. 4 *a*, (^{15}N , ^1H)-HMQC spectra of the ^{15}N -labelled POU₅ and the complex with duplex DNA, recorded with solvent presaturation. Shown are the regions containing the side chain amide resonances of Asn and Gln; in this region the N_H resonances of Arg 15 and 45 also are present (these resonances were folded in). Bottom: reference spectrum of 1 mM protein in 0.1 M NaCl, 5 mM dithiothreitol, $\text{H}_2\text{O}/\text{D}_2\text{O}=95/5$, pH 5.0 at 298 K, ^1H frequency of 500.13 MHz. The assignments for the side-chain resonances of the Asn, Gln and Arg residues are indicated. Middle: spectrum of the protein in the presence of 0.1 molar equivalent duplex GAATATGCAA (at 0.2 M NaCl). In the spectrum the assignments for resonances which are perturbed are indicated with dashed lines. Top: protein in the presence of 0.1 molar equivalent duplex at high-salt conditions (0.7 M NaCl). The resonances, which were broadened in the complex with DNA and which reappear under high-salt conditions, are indicated. The strongest broadening effects were observed

for Gln 23 and Arg 45. *b*, Model of the interaction of POU₅ with DNA. The view is along the helix III, which is positioned in the major groove. Shown are the positions of the sidechains of residues 23, 40, 45, 50 and 55 (indicated as yellow spheres), which all display broadening effects. The residues that do not display broadening are located at the surface of the protein away from the DNA binding site. *c*, Schematic model of the POU-DNA complex. The POU-homeodomain is docked onto B-DNA with the invariant amino-acid Asn 51 in the third homeodomain helix with respect to the invariant adenine in the DNA sequence (ATGCAAT)¹⁵. The N terminus of the POU-homeodomain is positioned in the minor groove. POU_s is modelled to enable recognition of the ATGC sequence of the octamer motif. The helices are drawn as cylinders. The linker sequence, connecting the POU_s with the POU-homeodomain, is drawn with a dashed line. Note that the DNA is drawn as regular B-DNA and that DNA bending⁷ has not been considered.

reappear again, presumably due to the dissociation of the complex. The resonances for the other Asn, Gln and Arg residues are unaffected under all conditions. Gln 40, Arg 45 and Asn 50 are located in and at the end of helix III, and are surface exposed. Inspection of the structure of POU₅ reveals that the residues Gln 23 and Asn 55 are located in the vicinity of helix III. This indicates that helix III of POU₅ and the surrounding regions are involved in the interaction with DNA. This helix corresponds with the recognition helix of λ -repressor. Helices II and III of POU₅ form a helix-turn-helix-related motif with a six-residue insertion after Gly 31.

To illustrate this point, POU₅ was docked onto DNA to fit helix III in the major groove¹³ (Fig. 4*b*). Residues 40, 45 and 50 are all in favourable positions to make direct interactions in the major groove, and residues 23 and 55 could make interactions with the DNA backbone. Other residues in helix III that are present at the protein-DNA interface are Thr 41, Ser 44, Glu 47, Ala 48 and Leu 49. The side chain of the conserved residue Leu 49 is located in the major groove of the DNA, and presum-

ably makes van der Waals interactions with one of the thymine methyl groups.

We also considered the mode of DNA binding by the complete POU domain. From NMR and crystallographic work the structures of several homeodomain-DNA complexes are known¹³⁻¹⁵. Considering the high sequence homology we assume a similar interaction of POU_{hd} with DNA^{9,15}. For the positioning of POU_s in the adjacent major groove to account for recognition of the ATGC sequence¹³, there are still two possible orientations, one with the C terminus 26 Å apart from the N terminus of POU_{hd}, and another with a distance of 41 Å. Because the linker between POU_s and POU_{hd} in a number of functional POU proteins is as short as 14 residues, this strongly favours the position with the shortest distance. This configuration is shown in Fig. 4*c*. An ethylation interference study of the POU-DNA complex¹⁶ showed one phosphate contact of the top strand between base-pair -2 and -1 preceding the octamer sequence, which cannot be accounted for by the protection of POU_{hd}. In our model this phosphate is in close contact with residues located

in the middle of helix III of POU_s. The model also accounts for the large unprotected region of the sugar-phosphate backbone as detected by hydroxyl-radical footprinting experiments¹⁶ near the middle of octamer binding site.

This work provides a framework for understanding protein-DNA interactions by proteins of the POU family and in particular their bipartite DNA-recognition. Future structural work on the complex and site-directed mutagenesis will be aimed at understanding the recognition specificity of the POU domain. □

Received 31 December 1992; accepted 8 April 1993.

- Herr, W. *et al. Genes Dev.* **2**, 1513-1516 (1988).
- Ruvkun, G. & Finney, M. *Cell* **64**, 475-478 (1991).
- Rosenfeld, M. G. *Genes Dev.* **5**, 897-907 (1991).
- Schöler, H. R. *Trends Genet.* **7**, 323-329 (1991).
- Schaffner, W. *Trends Genet.* **5**, 37-39 (1989).
- Prujn, G. J. M., van Driel, W. & van der Vliet, P. C. *Nature* **322**, 656-659 (1986).
- Verrijzer, C. P., Kal, A. J. & van der Vliet, P. C. *EMBO J.* **9**, 1883-1888 (1990).

- Botfield, M. C., Jancso, A. & Weiss, M. A. *Biochemistry* **31**, 5841-5848 (1992).
- Verrijzer, C. P. *et al. EMBO J.* **11**, 4993-5003 (1992).
- Havel, T. F. *Progr. Biophys. molec. Biol.* **56**, 43-78 (1991).
- O'Shea, E. K., Klemm, J. D., Kim, P. S. & Alber, T. *Science* **254**, 539-544 (1991).
- Harrison, S. C. *Nature* **353**, 715-719 (1991).
- Verrijzer, C. P., Kal, A. J. & van der Vliet, P. C. *Genes Dev.* **4**, 1964-1974 (1990).
- Kissinger, C. R., Liu, B., Martin-Blanco, E., Kornberg, T. B. & Pabo, C. O. *Cell* **63**, 579-590 (1990).
- Qian, Y. C. *et al. Cell* **59**, 573-580 (1989).
- Wolberger, C., Verstow, A. K., Lin, B., Johnson, A. & Pabo, C. O. *Cell* **67**, 517-528 (1991).
- Prujn, G. J. M., van Miltenburg, R. T., Claessens, J. A. J. & van der Vliet, P. C. *J. Virol.* **62**, 3092-3102 (1988).
- Tanaka, M. & Herr, W. *Cell* **60**, 375-389 (1990).
- Sturm, R. A., Das, G. & Herr, W. *Genes Dev.* **2**, 1582-1599 (1988).
- Janknecht, R. *et al. Proc. Natn. Acad. Sci. U.S.A.* **88**, 8972-8976 (1991).
- Kraulis, P. J. *Appl. Crystallogr.* **24**, 946-950 (1991).
- Wüthrich, K., Billeter, M. & Braun, W. *J. molec. biol.* **169**, 949-961 (1983).
- Koning, T. M. G., Boelens, R. & Kaptein, R. *J. magn. Reson.* **90**, 111-123 (1990).
- Kay, L. E. & Bax, A. *J. magn. Res.* **86**, 110-126 (1990).
- Zuiderweg, E. R. P., Boelens, R. & Kaptein, R. *Biopolymers* **24**, 601-611 (1985).
- Holm, L. & Sander, C. *J. molec. Biol.* (in the press).

ACKNOWLEDGEMENTS. We thank C. Sander and L. Holm (EMBL) for noting the structural homology with λ -repressor, using the program Dali²⁵. The investigations were supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO).

Rhodopsin phosphorylation as a mechanism of cyclic GMP phosphodiesterase regulation by S-modulin

Satoru Kawamura

Department of Physiology, Keio University School of Medicine, Shinano-machi 35, Shinjuku-ku, Tokyo 160, Japan

DURING light-adaptation by the vertebrate eye, the rods are desensitized and the light response is accelerated^{1,2}. When light is absorbed by the rods, a phosphodiesterase is activated that hydrolyses cyclic GMP^{3,4}. A light-induced decrease in cytoplasmic Ca²⁺ concentration⁵⁻⁷ is part of this light-adaptation process^{8,9}. The protein S-modulin (M_r 26,000) is known to increase the fraction of light-activated cyclic GMP-phosphodiesterase (PDE) at high Ca²⁺ concentrations in frog rod photoreceptors¹⁰. Here I present evidence that S-modulin lengthens the lifetime of active PDE (PDE*) at high Ca²⁺ concentrations. These S-modulin effects

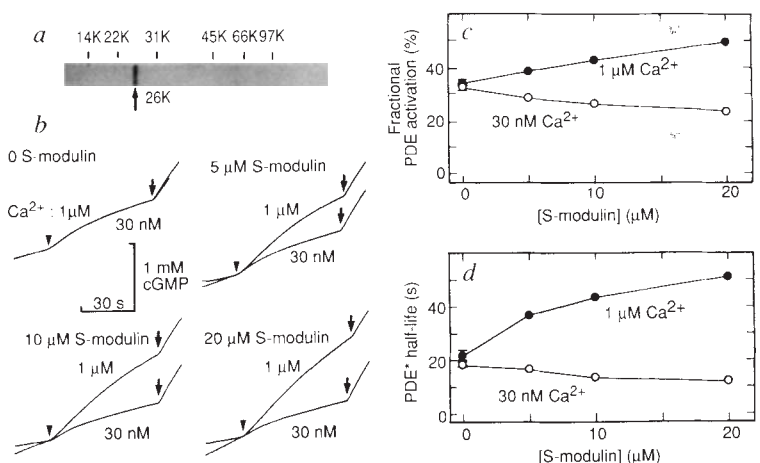
are observed in the physiological range of Ca²⁺ concentration (30 nM to 1 μ M; half-maximum effects at 200-400 nM). At the high Ca²⁺ concentrations at which S-modulin prolongs the lifetime of PDE*, S-modulin inhibits rhodopsin phosphorylation (half-maximum effect at ~100 nM Ca²⁺). ATP is necessary for the S-modulin effects on PDE activation. I therefore conclude that the Ca²⁺-dependent regulation of PDE by S-modulin is mediated by rhodopsin phosphorylation. This regulation seems to be the principal mechanism of light adaptation in vertebrate photoreceptors.

S-modulin was purified as a single band on an SDS-polyacrylamide gel (Fig. 1a)¹¹. The activity of S-modulin was tested using a pH assay at concentrations of 5-20 μ M S-modulin (Fig. 1b), which are higher than before (2 μ M in ref. 10) and close to that found in intact rods (about 40 μ M)¹⁰. The total concentration of rhodopsin (an integral membrane protein) was 6-8 μ M, because the concentration of disk membranes in the preparation was much lower than intact rods. However, as disk membranes were used, the rhodopsin concentration was that of intact rods at the site of reaction.

Measurements were made at two fixed Ca²⁺ concentrations (30 nM and 1 μ M). As reported previously¹⁰, in the absence of S-modulin (Fig. 1b; 0 S-modulin), PDE activation was not

FIG. 1 Dose-response of S-modulin on PDE activation. *a*, Coomassie brilliant blue staining of S-modulin (2 μ g) purified as described¹¹ to >98% as judged by densitometry after staining. *b*, PDE activation in the absence and presence of S-modulin. PDE activation was monitored by pH assay²⁰, giving a weak test flash that bleached 10⁴ rhodopsins per rod outer segment (ROS) (arrowhead) at 30 nM and 1 μ M Ca²⁺ in the absence and presence of 5-20 μ M S-modulin. The maximum PDE activity was measured by giving a saturating steady light that bleached 4 $\times$ 10⁷ rhodopsins per ROS per second (arrow). The concentration of S-modulin (S-mod) for each trial is indicated. *c*, Fractional PDE activation and *d*, PDE* half-life as functions of S-modulin concentration. Each value was obtained from the pH traces shown in *b*. Points at 0 μ M S-modulin are the mean $\pm$ s.d. of three experiments, including that in *b*.

METHODS. Frog (*Rana catesbeiana*) rod outer segments were isolated in K-gluconate buffer (115 mM potassium gluconate, 2.5 mM KCl, 2 mM MgCl₂, 1 mM dithiothreitol, 0.1 mM CaCl₂, 0.2 mM EGTA, 10 mM HEPES, pH 7.5) in complete darkness with the aid of an infrared converter. PDE activity was measured by pH assay²⁰. The pH drop induced by hydrolysis of cGMP²¹ was recorded and the PDE activity determined from the slope of the pH trace. To measure PDE activation, a weak test flash was first given in the presence of 0.5 mM ATP, 0.5 mM GTP, 4 mM cGMP and ROS membranes containing 6-8 μ M rhodopsin. Twenty to sixty seconds after the flash, a saturating steady light was given to measure the maximum PDE activity. Fractional PDE activation



was calculated from the ratio between the initial PDE activity elicited by the test flash and the maximum PDE activity. PDE* half-life was determined at the time when the slope of the pH trace reached half the initial slope of the pH change induced by the test flash. The deviation among the determined values was within 5% of the mean in several independent determinations of a single pH trace. To measure the light-induced changes in PDE activity, dark PDE activity was always subtracted.

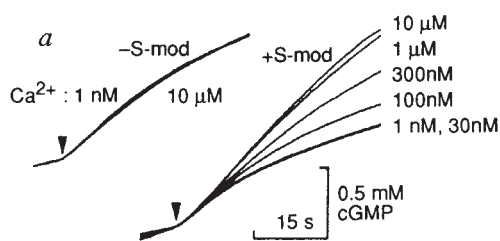
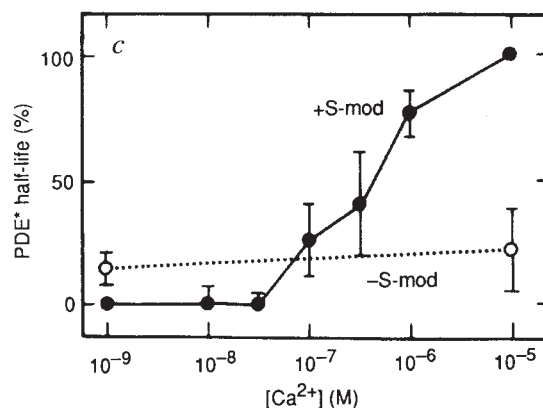
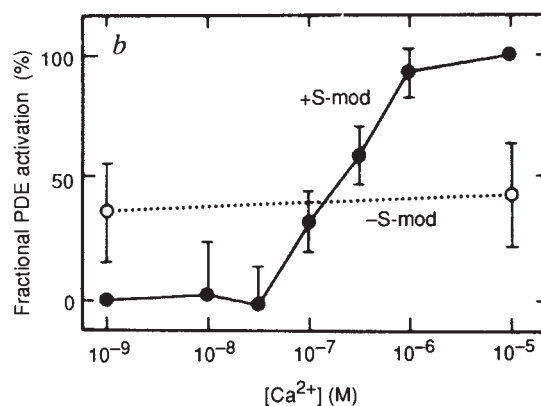


FIG. 2 Dependence of S-modulin effect on Ca^{2+} concentration. *a*, PDE activation measured at various Ca^{2+} concentrations. pH decrease was recorded in the absence and presence (8.7 μM) of S-modulin. Ca^{2+} concentration was buffered using a Ca^{2+} /EGTA buffering system²². Light stimulation was as for Fig. 1. The calculated values of fractional PDE activation in the records shown in the figure were 37–64%, and those of PDE* half-life were 18–39 s. These values are incorporated in the pooled data shown in *b* and *c*. *b*, Pooled data ($n=5$) of fractional PDE activation as a function of Ca^{2+} concentration (filled circles) in experiments where S-modulin concentration varied from 1.3 to 8.7 μM . Open circles indicate fractional activation without S-modulin. Because the extent of the S-modulin effect varied according to the concentration of S-modulin (Fig. 1*c* and *d*), in each experiment, fractional PDE activations at 1 nM Ca^{2+} and at 10 μM Ca^{2+} were arbitrarily set at 0 and 100%, respectively. Other data points were expressed as relative values on this scale. *c*, Pooled data ($n=3$) of PDE* half-life as a function of Ca^{2+} concentration in the presence of 3.2–8.7 μM S-modulin. In each experiment, the values at 1 nM Ca^{2+} and 10 μM Ca^{2+} were arbitrarily set at 0 and 100%, respectively. Open circles show PDE* half-life without S-modulin. Calculated Hill coefficients were 1.4 and 1.1 in *b* and *c*, respectively.



affected by Ca^{2+} either by a weak test flash (arrowhead) or by a saturating light (arrow). When S-modulin was added, PDE activation induced by the weak test flash was affected by the Ca^{2+} in three ways (Fig. 1*b–d*). First, in agreement with our previous study¹⁰, fractional PDE activation (ratio of the PDE activity induced by the weak test flash, to that induced by the saturating light) was higher at 1 μM Ca^{2+} than at 30 nM Ca^{2+} . Second, the time required for the recovery of PDE activity to the dark level was longer at 1 μM Ca^{2+} than at 30 nM Ca^{2+} , which indicated that PDE* lifetime is longer at high Ca^{2+} concentrations, in agreement with previous biochemical^{12,13} and electrophysiological¹⁴ results. Third, S-modulin had dual effects: it increased the fraction of activated PDE and prolonged its lifetime at 1 μM Ca^{2+} , but reduced both at 30 nM Ca^{2+} .

I measured PDE activation at various Ca^{2+} concentrations using a fixed concentration of S-modulin (8.7 μM) (Fig. 2*a*). Figure 2*b, c* summarizes the Ca^{2+} -dependency of the S-modulin effects that are evident at 30 nM–1 μM Ca^{2+} , with half-maximum effects at 200–400 nM Ca^{2+} . Because this range of Ca^{2+} concentration is physiological (~300 nM in the dark and ~140 nM in the light^{5–7}), it seems reasonable to conclude that Ca^{2+} -dependent regulation of PDE by S-modulin is the underlying mechanism of photoreceptor light adaptation.

Activation of PDE is terminated by phosphorylation of bleached rhodopsin with the γ -phosphate of ATP¹⁵. This could indicate that S-modulin lengthens PDE* lifetime at high Ca^{2+} concentrations by inhibiting rhodopsin phosphorylation, which is demonstrated by the following results.

Figure 3*a* is an autoradiogram of the incorporation of the γ -³²P of ATP into rhodopsin (arrowed). In the presence of S-modulin (+S-mod), rhodopsin phosphorylation is inhibited at high Ca^{2+} concentrations both in the dark and in the light. Radioactivity incorporated into rhodopsin indicates that the

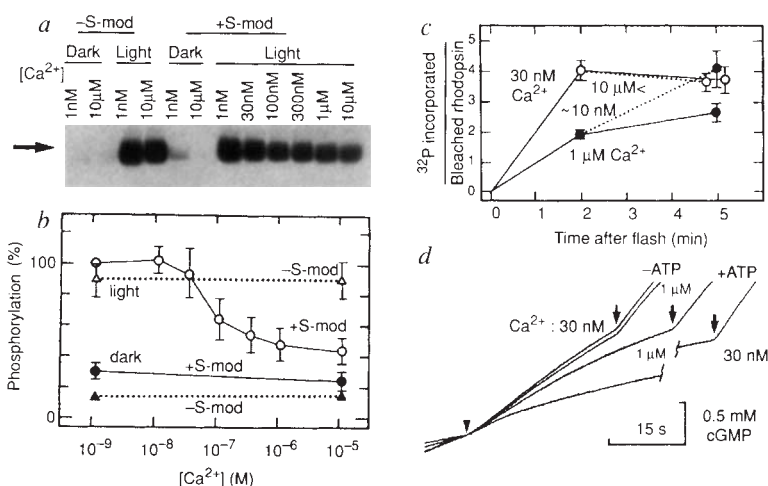
IC_{50} (concentration giving 50% inhibition) is ~100 nM Ca^{2+} (Fig. 3*b*). In the absence of S-modulin, however, rhodopsin phosphorylation is almost unaffected by Ca^{2+} , either in the dark or in the light (Fig. 3*a* and *b*).

To show that the decreased phosphorylation at high Ca^{2+} concentrations is due to suppression of kinase activity rather than the activation of phosphatase, I measured rhodopsin phosphorylation at constant Ca^{2+} concentration (Fig. 3*c*, solid line), or with variable Ca^{2+} concentrations during the incubation (dotted line). If a Ca^{2+} -activated phosphatase activity is present, phosphorylation should be reduced when the Ca^{2+} concentration is raised; however, the level did not change even when the Ca^{2+} concentration was increased from 30 nM to >10 μM (circles). It is unclear, however, whether S-modulin inhibits the kinase itself or prevents interaction between the enzyme and bleached rhodopsin.

As rhodopsin phosphorylation is the principal reaction of ATP in the phototransduction cascade^{3,4}, PDE activation was measured in the presence of S-modulin but in the absence of ATP to investigate the role of rhodopsin phosphorylation in S-modulin action. Neither fractional PDE activation nor PDE* lifetime was affected by Ca^{2+} when ATP was omitted (Fig. 3*d* and legend). I therefore conclude that S-modulin action is mediated through rhodopsin phosphorylation.

These results indicate that the action of S-modulin is more complicated than previously suggested¹⁰. A large effect on PDE* lifetime has been detected (Fig. 1), which could be explained by assuming that the purified S-modulin used here is not a single species but a mixture of isoforms. ATP has two distinct effects on PDE activation¹⁶: it reduces PDE* lifetime (half-effects at ~2 μM) and reduces fractional PDE activation (at ~50 μM). It is possible that only the latter was inhibited by a single form of S-modulin in the previous study, but that both are inhibited

FIG. 3 Effect of S-modulin on rhodopsin phosphorylation. *a*, ^{32}P -autoradiograph showing rhodopsin phosphorylation under various conditions. When present, S-modulin concentration was 10 μM . *b*, Pooled data ($n=4$) obtained by counting ^{32}P rhodopsin bands dissected from the SDS-polyacrylamide gel. Each point represents the mean $\pm$ s.d. Each phosphorylation level was normalized to that obtained in the light at 1 nM Ca^{2+} with S-modulin. Closed and open symbols show the phosphorylation levels in the dark and light, respectively. Dark phosphorylation without S-modulin (filled triangles) was examined only once. *c*, Indication that rhodopsin kinase activity is inhibited at high Ca^{2+} concentrations in the presence of 10 μM S-modulin. After a light flash, rhodopsin was phosphorylated at high (1 μM) and low (30 nM) Ca^{2+} concentrations. At each Ca^{2+} concentration, 2 min after the light flash the Ca^{2+} concentration in one portion was changed (dotted line; from 30 nM to $>10 \mu\text{M}$ or from 1 μM to $\sim 10 \text{nM}$), while the Ca^{2+} concentration in the other portion was maintained at the same level (solid line). The radioactivity of ^{32}P in the dissected rhodopsin bands of the SDS gel was counted and the number of ^{32}P incorporated into a bleached rhodopsin was calculated. Each data point is the mean $\pm$ s.d. of three independent experiments. *d*, ATP effect on PDE activation examined in the presence of 8.7 μM S-modulin. PDE activation was monitored at 30 nM and 1 μM Ca^{2+} in the absence and presence (500 μM) of ATP. Light stimulation was the same as in Fig. 1. The calculated values of fractional PDE activation are: without ATP, 38.5% (30 nM Ca^{2+}) and 35.7% (1 μM); with ATP, 23.0% (30 nM) and 35.7% (1 μM). PDE* half-life values are: without ATP, 27.0 s (30 nM Ca^{2+}) and 27.1 s (1 μM); with ATP, 9.6 s (30 nM) and 25.9 s (1 μM). The interval in the lowest record was 30 s. The intensity of the light flash used in the phosphorylation experiment was 350 times higher than that used for the measurement of PDE activation (see Fig. 1 legend, and below). As the PDE* half-life is dependent on the intensity of light flash¹⁶, direct comparison between the phosphorylation experiment and the PDE activity measurement is not possible.



METHODS. A ROS membrane suspension (15 μl) containing 10 μg rhodopsin was mixed with 10 μl of a [γ - ^{32}P]ATP solution (24 MBq μmol^{-1} of ATP) in the dark in a glass tube. The resultant reaction mixture (25 μl) contained (final concentration): rhodopsin (10 μM), ATP (100 μM), GTP (500 μM), cGMP (4 mM). Thirty seconds after mixing, a light flash bleaching 3.5×10^6 rhodopsins per ROS was given. The reaction was terminated 2 min after the flash by addition of 200 μl 10% trichloroacetic acid. The sample was then centrifuged (10,000g for 10 min) and the precipitate washed once with 1 ml of potassium-gluconate buffer. After centrifugation, the precipitate was subjected to SDS-PAGE²³. The gel was stained with Coomassie brilliant blue and autoradiographed; rhodopsin bands were dissected out and ^{32}P estimated by Cerenkov counting in a plastic tube.

by a mixture of isoforms of S-modulin in this study, perhaps as a result of the different purification methods used.

Recoverin is an S-modulin-like 26K Ca^{2+} -binding protein from bovine rods which modulates guanylyl cyclase activity^{17,18}. As recoverin shows a Ca^{2+} -dependent inhibition of rhodopsin phosphorylation (K. Palczewski, personal communication, and S.K., unpublished observation), it is likely that recoverin is the same protein as S-modulin.

Phosphorylation of receptor molecules may be a general mechanism for regulating signal transduction¹⁹. It will be interesting to see whether regulation mechanisms similar to the one described here are used in other transduction systems. □

Oncoprotein MDM2 conceals the activation domain of tumour suppressor p53

Jonathan D. Oliner^{*†}, Jennifer A. Pietenpol^{*}, Sam Thiagalingam^{*}, Jeno Gyuris[‡], Kenneth W. Kinzler^{*} & Bert Vogelstein^{*†§}

^{*} The Johns Hopkins Oncology Centre, 424 North Bond Street, Baltimore, Maryland 21231, USA

[†] Program in Human Genetics, The Johns Hopkins University, School of Medicine, Baltimore, Maryland 21231, USA

[‡] Department of Genetics, Harvard Medical School, and Department of Molecular Biology, Massachusetts General Hospital, Boston, Massachusetts 02114, USA

THE tumour-suppressor gene *p53* is inactivated in most human malignancies¹ either by missense mutations¹ or by binding to oncogenic proteins²⁻⁴. In human soft tissue sarcomas, inactivation apparently results from *MDM2* gene amplification⁴. *MDM2* is an oncogene product^{5,6} that may function by binding to *p53* and inhibiting its ability to activate transcription³. Here we show that, when expressed in *Saccharomyces cerevisiae*, human *MDM2* inhibits human *p53*'s ability to stimulate transcription by binding to a region that nearly coincides with the *p53* acidic activation domain. The isolated *p53* activation domain fused to another DNA-binding protein is also inactivated by *MDM2*, confirming that *MDM2* can inhibit *p53* function by concealing the activation domain of *p53* from the cellular transcription machinery.

To determine whether *MDM2* affects transcriptional activation by *p53* in this system, expression vectors encoding both proteins were transfected into yeast, together with a

Received 10 August 1992; accepted 5 February 1993.

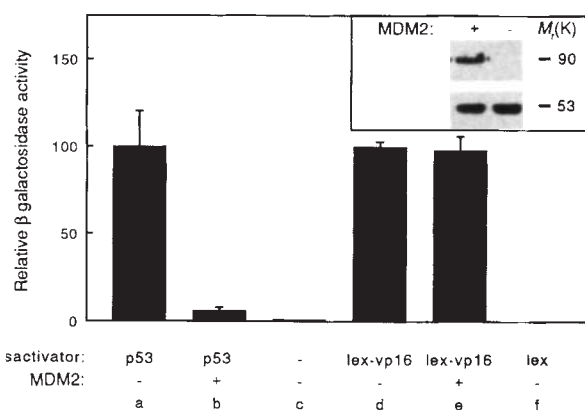
- Baylor, D. A. & Hodgkin, A. L. *J. Physiol.* **242**, 729-758 (1974).
- Baylor, D. A., Lamb, T. D. & Yau, K.-W. *J. Physiol.* **288**, 589-611 (1979).
- Stryer, L. *A. Rev. Neurosci.* **9**, 87-119 (1986).
- Kaupp, U. B. & Koch, K.-W. *A. Rev. Physiol.* **54**, 153-175 (1992).
- Yau, K.-W. & Nakatani, K. *Nature* **313**, 579-582 (1985).
- McNaughton, P. A., Cervetto, L. & Nunn, B. J. *Nature* **322**, 261-263 (1986).
- Ratto, G. M., Payne, R., Owen, W. G. & Tsien, R. Y. *J. Neurosci.* **8**, 3240-3246 (1988).
- Matthews, H. R., Murphy, R. L. W., Fain, G. & Lam, T. D. *Nature* **334**, 67-69 (1988).
- Nakatani, K. & Yau, K.-W. *Nature* **334**, 69-71 (1988).
- Kawamura, S. & Murakami, M. *Nature* **349**, 420-423 (1991).
- Kawamura, S., Takamatsu, K. & Kitamura, K. *Biochem. biophys. Res. Commun.* **186**, 411-417 (1992).
- Kawamura, S. & Bownds, M. D. *J. gen. Physiol.* **77**, 571-591 (1981).
- Barkdoll, A. E. III, Pugh, E. N. Jr & Sitaramayya, A. *J. gen. Physiol.* **93**, 1091-1108 (1989).
- Torre, V., Matthews, H. R. & Lamb, T. D. *Proc. natn. Acad. Sci. U.S.A.* **83**, 7109-7113 (1986).
- Liebman, P. A. & Pugh, E. N. Jr *Nature* **287**, 734-736 (1980).
- Kawamura, S. *Biochim. biophys. Acta* **732**, 276-281 (1983).
- Dizhoor, A. M. et al. *Science* **251**, 915-918 (1991).
- Lambrecht, H.-G. & Koch, K.-W. *EMBO J.* **10**, 793-798 (1991).
- Benovic, J. L., Bouvier, M., Caron, M. G. & Lefkowitz, R. J. *A. Rev. Cell Biol.* **4**, 405-428 (1988).
- Kawamura, S. & Murakami, M. *Biochim. biophys. Acta* **870**, 256-266 (1986).
- Yee, R. & Liebman, P. A. *J. biol. Chem.* **253**, 8902-8909 (1978).
- Kawamura, S. & Murakami, M. *J. gen. Physiol.* **94**, 649-668 (1989).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).

ACKNOWLEDGEMENTS. I thank A. Kaneko, S. Watanabe and G. Ayoub for discussion. This work was supported by grants from the Ministry of Education, Science and Culture of Japan, Nissan Science Foundation and the Naito Foundation (to S.K.), and the Human Frontier Science Program Organization (to A. Kaneko).

§ To whom correspondence should be addressed.

FIG. 1 Inhibition of p53-mediated transactivation by MDM2. Yeasts were stably transfected with expression plasmids encoding p53, lex-VP16, MDM2 or the appropriate vector-only controls, as indicated. p53-responsive (bars a-c) or lexA-responsive (bars d-f) β -galactosidase reporter plasmids were used to assess the response. Inset: Western blot analysis⁴ demonstrating MDM2 (90K) and p53 (53K) expression in representative yeast strains. The strain indicated by a plus sign was transfected with expression vectors encoding full-length MDM2 and p53; the strain indicated by a minus was transfected with only the p53 expression vector.

METHODS. The MDM2 expression plasmid, pPGK-MDM2, was constructed by inserting full-length MDM2 cDNA⁴ into pPGK²². Galactose-inducible p53 (pRS314SN²⁰), lexA-VP16 (YVlexA²³), and lexA (YVlexA minus VP16) plasmids were used as indicated. The reporters were PG16-*lacZ*⁸ (p53-responsive) and pJK 103^{9,27} (lexA-responsive). All plasmids were introduced into *S. cerevisiae* strain EGY48; ref. 25. Yeast strains represented by bars a-c were grown at 30 °C to mid-log phase in selective liquid medium containing 2% raffinose, induced for 30 min by addition of 2% galactose, collected, and lysed⁶. Strains represented by bars d-f were treated similarly, except that cells were induced in galactose for 4 h to obtain measurable levels of β -galactosidase. β -galactosidase activities of bars b and c are shown relative to that of bar a; β -galactosidase activities of bars e and f



are shown relative to that of bar d. For western blotting, anti-p53 antibody PAb1801 (53 K bands; Oncogene Science) or MDM2 polyclonal antibodies⁴ (90 K bands) were used.

p53-responsive reporter consisting of a β -galactosidase gene driven by a minimal promoter and a multimerized human DNA sequence which bound p53 strongly *in vitro*^{7,8}. Reporter expression was completely dependent on p53 (Fig. 1, a and c). MDM2 expression inhibited p53-mediated transactivation of this reporter 16-fold (Fig. 1, a and b) without affecting p53 expression (Fig. 1, inset).

To ensure that MDM2 was not downregulating transcription in general, a different transcriptional activator (lexA-VP16, consisting of the lexA DNA-binding domain fused to the VP16 acidic activation domain) and reporter (lexA-responsive site upstream of a β -galactosidase gene) were transfected into yeast together with the same MDM2 expression vector. LexA-VP16 transactivation was unaffected by the presence of MDM2 (Fig. 1, d and e), and MDM2 expression had no effect on the growth rate of the cells.

To map the domains of MDM2 and p53 responsible for their interaction, we transfected yeast with expression vectors encoding random fragments of MDM2 fused to the lexA DNA-binding domain (amino acids 2–202), together with an expression vector encoding full-length p53 (supplying its own activation domain) and a reporter containing a lexA-responsive site upstream of a β -galactosidase gene⁹. p53 binding by the MDM2 fragment could thus be detected by activation of the reporter²⁶. In an initial control experiment, full-length MDM2 fused to the lexA DNA-binding domain indeed activated the lexA-responsive reporter in the presence of p53, whereas the same expression constructs lacking the MDM2 or p53 cDNA inserts did not (Table 1, strains 1, 2 and 3). Of ~5,000 clones containing MDM2 fragments fused to the lexA DNA-binding domain, only 17 activated the reporter. The clones could be placed into two classes. The first class (two clones) expressed about 5-fold less β -galactosidase than the other fifteen clones and expression was independent of p53 (Fig. 2a). These two clones encoded MDM2 residues 190–340 and 269–379, respectively, so that both overlap the only acidic domain in MDM2 (residues 230–301). Of the residues in this domain, 37% are aspartic or glutamic acids, and none is basic. The domain seems to activate transcription only when isolated from the rest of the MDM2 sequence, as the entire MDM2 protein fused to lexA failed to activate the reporter (Table 1, strain 3). The other class (15 clones) each contained the amino-terminal region of MDM2 (Fig. 2b), and reporter activity required p53 co-expression. To narrow down the region of interaction, we made six smaller clones. The smallest tested region of MDM2 that could interact with full-length p53 contained codons 1 to 118 (Fig. 2b). The relatively large size of this region is consistent with the absence of reporter activation when smaller fragments of MDM2 were used in the assay (200 base pairs (bp) instead of 600 bp average size).

In the converse set of experiments, yeast clones containing p53 fragments fused to the B42 acidic activation domain were screened for lexA-responsive reporter expression in the presence of the lexA DNA-binding domain fused to full-length MDM2. The 14 positive clones obtained fell into three subsets, containing residues 1–41, 13–57, and 1–50, respectively (Fig. 2c). The minimal overlap between these fragments contains codons 13–41, but a further 9–12 amino acids on either side of these residues were apparently necessary for activity (Fig. 2c). To confirm that the amino-terminal region of p53 was required for MDM2 binding, we generated an additional yeast strain expressing the lexA DNA-binding domain fused to residues 74–393 of p53 and the B42 acidic activation domain fused to full-length MDM2. Although the proteins were expressed at the expected levels, as in all the other strains described here (examples in Fig. 2d, e), the lexA-responsive reporter was not activated (Table 1, strain 8), showing that the amino terminus of p53 is required for the interaction. To test whether this held true *in vitro*, chemically synthesized peptides containing the amino-terminal portion of

TABLE 1 Summary of representative results

Strain number	p53 construct	MDM2 construct	Activation§
1	p53*	Vector†	—
2	p53*	lexA-MDM2‡	+
3	Vector*	lexA-MDM2‡	—
4	p53*	lexA-MDM2 (1–118)‡	+
5	Vector*	lexA-MDM2 (1–118)‡	—
6	B42-p53 (1–41)‡	lexA-MDM2‡	+
7	B42-p53 (1–41)‡	Vector†	—
8	lexA-p53 (74–393)†	B42-MDM2‡	—
9	p53 (1–137)*	lexA-MDM2‡	—

The MDM2 and p53 proteins expressed in each strain are indicated, with the relevant reporters. Numbers in parentheses refer to the MDM2 or p53 amino acids encoded (absence of parentheses indicates full-length protein, that is, MDM2 amino acids 1 to 491 or p53 amino acids 1 to 393). Constructs are detailed in the legends to Figs 1, 2 and 3. The lexA-responsive β -galactosidase reporter plasmid (pJK 103; ref 19) was present in all strains, and the expression of all the constructs was verified by western blotting (examples in Figs 2 and 3).

* pRS314 vector²⁰.

† plex(1–202) + PL vector, containing lexA DNA-binding domain fused to insert²¹.

‡ pJG4–5 vector, containing B42 activation domain fused to insert (J.G., manuscript submitted).

§ A plus sign indicates that colonies turned blue after 24 h incubation on X-gal-containing selective medium; a minus sign indicates that colonies remained white after 72 h incubation.

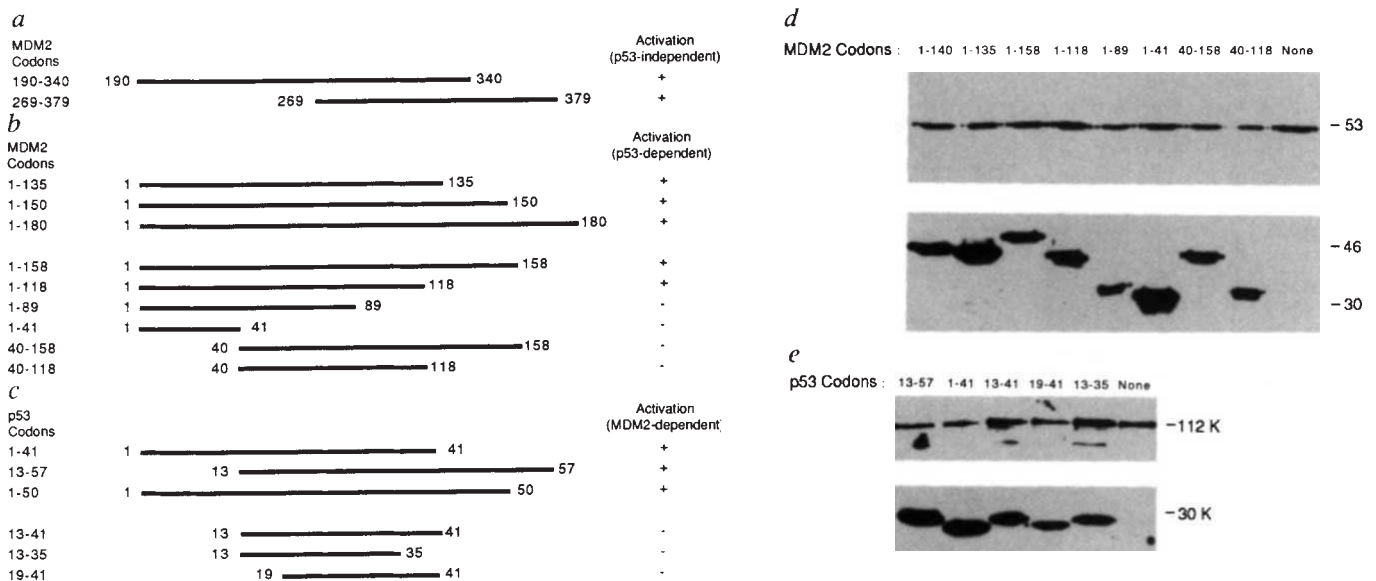


FIG. 2 Determination of MDM2 and p53 domains of interaction. **a** and **b**, Random fragments of MDM2 were fused to sequences encoding the lexA DNA-binding domain and the resultant clones transfected into yeast carrying pRS314SN (p53 expression vector) and pJK 103 (lexA-responsive β -galactosidase reporter). Yeast clones expressing β -galactosidase were identified by their blue colour, and the MDM2 sequences in the lexA fusion vector were determined. β -galactosidase activity was observed independent of p53 expression in **a**, but was dependent on p53 expression in **b**. The bottom 6 clones in **b** were generated by genetic engineering. **c**, Random fragments of p53 were fused to the sequence encoding the B42 acidic activation domain and a haemagglutinin epitope tag; the resultant clones were transfected into yeast carrying lexA-MDM2 (lex DNA-binding domain fused to full-length MDM2) and pJK103. Yeast clones were identified as described above and all were MDM2-dependent. The bottom three clones were generated by genetic engineering. **d**, **e**, Protein expression from the yeast strains in **a**–**c**. For western blotting 20 μ g protein was used per lane. The MDM2 and p53 codons contained in the fusion vectors are shown at the top of **d** and **e**, respectively. Upper panel in **d** was probed with PAb1801 detecting p53; the lower panel was probed with anti-lexA polyclonal antibodies (lex Ab) detecting MDM2 fusion proteins of M_r 30–50K. Upper panel in **e** was probed with lex Ab detecting the lexA full-length MDM2 fusion protein of 112K; the lower panel was probed with a monoclonal antibody directed against the haemagglutinin epitope tag (Berkeley Antibody), detecting p53 fusion proteins of ~25–30K.

METHODS. **a**, The insert from a full-length MDM2 cDNA clone was sheared into random fragments by sonication, ligated to linkers, and amplified by the

polymerase chain reaction²⁴. Fragments were fractionated on an acrylamide gel into size ranges of 100–400 bp or 400–1,000 bp, cloned into lexA(1–202) + PL (ref. 21), and transfected into bacteria (XL-1 Blue; Stratagene). At least 10,000 bacterial colonies were scraped off agar plates, and the plasmid DNA was transfected into a strain of EGY48 containing pRS314SN (p53 expression vector) and pJK103 (lexA-responsive β -galactosidase reporter). About 5,000 yeast clones were plated on selective medium containing 2% dextrose, and replica-plated onto galactose- and X-gal-containing selective medium. Blue colonies appeared only on plates with larger fragments of MDM2. The 17 isolated colonies were tested for blue colour in this assay in the presence and in the absence of galactose induction of p53. Plasmid DNA extracted from yeast clones was selectively transferred to bacterial strain KC8 as described (J.G., manuscript submitted). MDM2 sequences of the 15 p53-dependent clones coded for peptides ranging from 135 to 265 amino acids in length and began exclusively at the initiator methionine. Three of the MDM2 sequences are shown at the top of **b**. The lower six sequences were genetically engineered and subsequently tested to narrow the binding region further. **c**, Methods are essentially the same as those described in **a** and **b**, except that p53 fragments were cloned into pJG4–5 (J.G., manuscript submitted), producing a fusion protein C-terminal to the B42 acidic activation domain and incorporating an epitope of haemagglutinin. Clones were transfected into a strain of EGY48 containing lexA-MDM2 (lexA-202 + PL containing full-length MDM2) and pJK103. The top three p53 sequences were derived from yeast by colony screening; the lower three were genetically engineered to contain the indicated fragments.

p53 were incubated with ³⁵S-methionine-labelled p53 and MDM2 proteins generated in reticulocyte lysates, before precipitating the resulting p53–MDM2 complexes with antiserum against MDM2 (as described in ref. 4). As expected, peptides containing residues 1–41 or 6–41 inhibited p53–MDM2 complex formation by over 80% at a peptide concentration of 7×10^{-6} M, whereas peptides containing residues 16–41 or 340–358 inhibited it by less than 5% at the same concentration.

Strikingly, the region of p53 required to bind MDM2 was almost identical to the previously identified acidic activation domain of p53 (residues 20–42)^{10–13}, suggesting that MDM2 inhibits p53-mediated transcriptional activation by 'concealing' the activation domain of p53 from the transcriptional machinery. If so, the p53 activation domain by itself should still be inhibited by full-length MDM2. We therefore produced a protein containing the p53 activation domain (codons 1–73) fused to the lexA DNA-binding domain. As predicted from earlier work^{10,11}, this construct strongly activated the lexA-responsive reporter (Fig. 3a). Co-expression of MDM2 resulted in a fivefold decrease in reporter activity, showing that MDM2 can specifically inhibit the function of the p53 activation domain regardless of protein sequences tethering it to DNA (Fig. 3a).

If MDM2 binds to the P53 activation domain (Fig. 2) and

conceals it from the transcriptional machinery (Fig. 3a), how then can the lexA-MDM2–p53 complex activate transcription from the lexA-responsive reporter (Table 1, strain 2)? The only functional activation domain in the lexA-MDM2–p53 complex of strain 2 is that of p53, and it might therefore be predicted that it would be concealed by binding to MDM2. p53 exists as a homotetramer^{14–16}, however, and activation by the lexA-MDM2–p53 complex could thus be due to the presence of four activation domains, not all concealed by MDM2. To test this, the C-terminal domain of p53 required for homo-oligomerization^{15,17} was removed from the p53 expression construct, so that it consisted of only codons 1–137 (Table 1, strain 9) but retained the entire activation domain (as shown in Fig. 3a, bar a) and the entire domain required for interaction with MDM2 (Table 1, strain 6). This truncated p53 protein was unable to activate the lexA-responsive reporter by interacting with lexA-MDM2, even though binding to the lexA reporter was preserved (Fig. 3b, lane 2).

Although we cannot rule out the possibility that MDM2 also inhibits the ability of p53 to bind to specific DNA sequences, these results support a model in which MDM2 directly controls p53-regulated gene expression by concealing the activation domain of p53. GAL80 may inhibit GAL4 function in yeast by

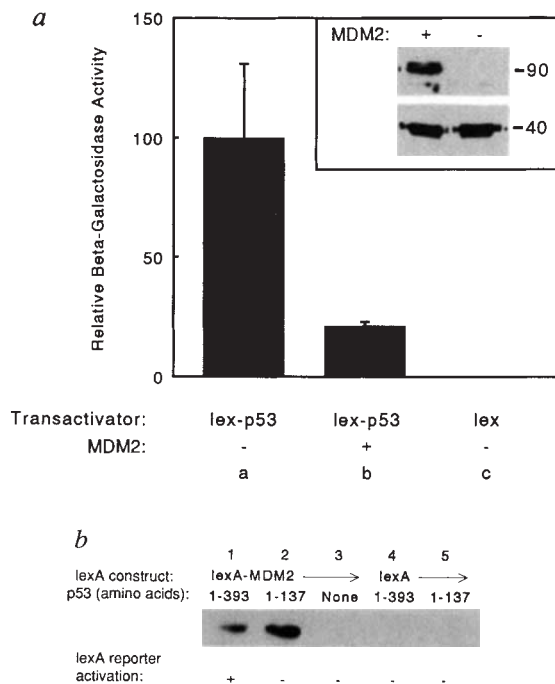


FIG. 3 Inhibition of the p53 activation domain by MDM2. *a*, Yeasts were stably transfected with expression vectors encoding a lexA-p53 (p53 codons 1–73) fusion (bars *a* and *b*) or lexA alone (bar *c*). Strain *b* also expressed full-length MDM2, and all strains contained the lexA-responsive β -galactosidase reporter plasmid. *b*, Lysates from yeast stably transfected with the indicated lexA and p53 expression vectors were assessed for lexA operator binding (top) and lexA reporter activation (bottom). For binding, extracts were incubated with a radiolabelled DNA fragment containing the lexA binding site²⁵, and then with an antibody against p53 to precipitate the lexA-MDM2-p53 complex. The precipitated DNA was purified and analysed by gel electrophoresis; an autoradiograph of the gel (bands in the middle row) shows that the 36-bp fragment was precipitated in both lanes 1 and 2.

METHODS. *a*, Essentially as described in Fig. 1 legend, except that the strains were grown to mid-log phase in 2% dextrose before galactose induction of p53 for 2 h. The lex-p53 construct was identical to lex-VP16 (YVlexA²³), except that VP16 sequences were replaced by p53 sequences encoding amino acids 1 to 73. The MDM2 and lex constructs used were the same as those described in Fig. 1. Inset: Upper panel probed with MDM2 polyclonal antibodies detecting full-length MDM2 (90K); lower panel probed with lex Ab detecting the lex-p53 fusion protein of 40K. *b*, Lanes 1, 3 and 4 represent the yeast described in Table 1 (strains 2, 3 and 1, respectively). Yeast strains used in lanes 2 and 5 differ from those used in lanes 1 and 4, respectively, only in that the p53 expression construct in lanes 2 and 5 encodes a carboxy-truncated version of p53 (amino acids 1–137). Yeast strains were grown as described for Fig. 1 and galactose-induced for 8 h. Lysates were tested for β -galactosidase activity and binding to a 36-bp DNA fragment containing the lexA DNA-binding site. DNA binding was assayed as described⁸ using p53 Ab1801 for immunoprecipitation.

a similar mechanism^{18,19}. It will be interesting to determine whether the activity of mammalian transcription factors is commonly regulated by concealing their activation domains. This mechanism suggests that tumours expressing abnormally high levels of MDM2 could be treated by means of peptides derived from the p53 interaction domain. By binding to MDM2 and preventing its interaction with p53, such peptides could restore p53 function to the tumour cells, presumably preventing further growth. □

- Oliner, J. D., Kinzler, K. W., Meltzer, P. S., George, D. L. & Vogelstein, B. *Nature* **358**, 80–83 (1992).
- Fakhrazadeh, S., Trusko, R. S. & George, D. *EMBO J.* **10**, 1565–1569 (1991).
- Finlay, C. A. *Molec. cell. Biol.* **13**, 301–306 (1993).
- Vogelstein, B. & Kinzler, K. W. *Cell* **70**, 523–526 (1992).
- Kern, S. E. *et al. Science* **256**, 827–830 (1992).
- Kamins, J., Richardson, P., Mosialos, G., Brent, R. & Gilmore, T. *Molec. cell. Biol.* **10**, 2840–2847 (1990).
- Fields, S. & Jang, S. K. *Science* **249**, 1046–1049 (1990).
- Raycroft, L., Wu, H. & Lozano, G. *Science* **249**, 1049–1051 (1990).
- Unger, T., Nau, M. M., Segal, S. & Minna, J. D. *EMBO J.* **11**, 1383–1390 (1992).
- Miller, C. W., Imai, Y., Aslo, A., Li, L. & Koefler, H. P. *Proc. Am. Assoc. Cancer Res.* **33**, 386 (1992).
- Stenger, J. E., Mayr, G. A., Mann, K. & Tegtmeyer, P. *Molec. Carcinogen.* **5**, 102–106 (1992).
- Sturzbecher, H. W. *et al. Oncogene* **7**, 1513–1523 (1992).
- Friedman, P. N., Chen, X., Bargonetti, J. & Prives, C. *Proc. natn. Acad. Sci. U.S.A.* (in the press).
- Milner, J. & Medcalf, E. A. *Cell* **65**, 765–774 (1991).
- Leuther, K. K. & Johnston, S. A. *Science* **256**, 1333–1335 (1992).
- Ma, J. & Ptashne, M. *Cell* **50**, 137–142 (1987).
- Nigro, J. M., Sikorski, R., Reed, S. I. & Vogelstein, B. *Molec. cell. Biol.* **12**, 1357–1365 (1992).
- Ruden, D. M., Ma, J., Li, Y., Wood, K. & Ptashne, M. *Nature* **350**, 250–252 (1991).
- Poon, D. *et al. Molec. cell. Biol.* **11**, 4809–4821 (1991).
- Dalton, S. & Treisman, R. *Cell* **68**, 597–612 (1992).
- Kinzler, K. W. & Vogelstein, B. *Nucleic Acids Res.* **17**, 3645–3653 (1989).
- Golemis, E. A. & Brent, R. *Molec. cell. Biol.* **12**, 3006–3014 (1992).
- Fields, S. & Song, O. *Nature* **340**, 245–246 (1989).
- Ebina, Y. Y. *et al. J. Biol. Chem.* **258**, 13258–13261 (1983).

ACKNOWLEDGEMENTS. We thank R. Brent for unpublished material, discussion and support, E. Golemis for EGY48, S. Dalton and R. Treisman for YVlexA, P. A. Weil for pPGK, E. Heimer for oligopeptides, and T. Gwiazda for preparing the manuscript. This work was supported in part by grants from the Clayton Fund, The National Foundation for Cancer Research, the Human Frontiers Science Program (to J.G.) and from the NIH.

Yeast DNA repair and recombination proteins Rad1 and Rad10 constitute a single-stranded-DNA endonuclease

Alan E. Tomkinson*, A. Jane Bardwell, Lee Bardwell, Nancy J. Tappe & Errol C. Friedberg

Laboratory of Molecular Pathology, Department of Pathology, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Boulevard, Dallas, Texas 75235-9072, USA

DAMAGE-SPECIFIC recognition and incision of DNA during nucleotide excision repair in yeast¹ and mammalian cells² requires multiple gene products. Amino-acid sequence homology between several yeast and mammalian genes suggests that the mechanism of nucleotide excision repair is conserved in eukaryotes^{2–7}, but very little is known about its biochemistry. In the yeast *Saccharomyces cerevisiae* at least 6 genes are needed for this process, including *RAD1* and *RAD10* (ref. 1). Mutations in the two genes inactivate nucleotide excision repair^{8,9} and result in a reduced efficiency of mitotic recombinational events between repeated sequences^{10–15}. The Rad10 protein has a stable and specific interaction with Rad1 protein^{16,17} and also binds to single-stranded DNA and promotes annealing of homologous single-stranded DNA¹⁸. The amino-acid sequence of the yeast Rad10 protein is homologous with that of the human excision repair gene *ERCC1* (ref. 3). Here we demonstrate that a complex of purified Rad1 and Rad10 proteins specifically degrades single-stranded DNA by an endonucleolytic mechanism. This endonuclease activity is presumably required to remove non-homologous regions of single-stranded DNA during mitotic recombination between repeated sequences as previously suggested¹³, and may also be responsible for the specific incision of damaged DNA during nucleotide excision repair.

We have purified Rad1 protein to >85% homogeneity (Fig. 1). The purified protein has a relative molecular mass of ~150,000 as determined by gel electrophoresis. Similar estimates were obtained following overexpression of Rad1 protein in

Received 20 November 1992; accepted 26 February 1993.

- Hollstein, M., Sidransky, D., Vogelstein, B. & Harris, C. C. *Science* **253**, 49–53 (1991).
- Scheffner, M., Werness, B. A., Huibregtse, J. M., Levine, A. J. & Howley, P. M. *Cell* **63**, 1129–1136 (1990).
- Momand, J., Zambetti, G. P., Olson, D. C., George, D. & Levine, A. J. *Cell* **69**, 1237–1245 (1992).

* Present address: Institute of Biotechnology, Center for Molecular Medicine, University of Texas Health Science Center at San Antonio, 15355 Lambda Drive, San Antonio, Texas 78245-3207, USA.

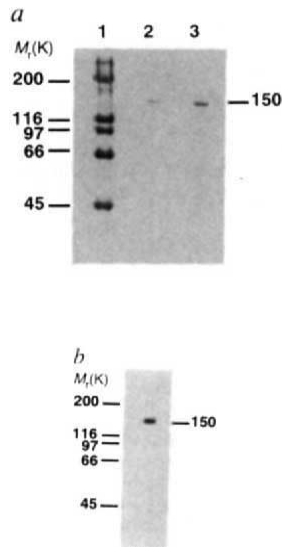


FIG. 1 Purification of Rad1 protein. *a*, Purified Rad1 protein was electrophoresed through a 7.5% denaturing polyacrylamide gel and stained with Coomassie blue. Lane 1, M_r standards (BioRad); lane 2, 0.35 μ g Rad1 protein; lane 3, 0.7 μ g Rad1 protein. *b*, After electrophoresis, Rad1 protein (0.14 μ g) was transferred to nitrocellulose which was incubated with affinity-purified antibodies raised against Rad1 expressed in *E. coli* (A.J.B. and E.C.F., unpublished results). Immune complexes were detected by enhanced chemiluminescence (Amersham). The apparent M_r of Rad1 protein ($\sim$ 150K) is indicated.

METHODS. Purification of Rad1 protein was achieved by overexpression of the protein from the cloned *RAD1* gene in a yeast strain deleted of the *RAD10* gene²⁷. Details of the purification will be presented elsewhere (A.J.B., L.B., E.C.F., N.J.T. and A.E.T., manuscript in preparation). Briefly, a cleared lysate was prepared from 77 g (wet weight) of cells from strain BJ *RAD10* Δ (pG12-*RAD1*) as described²⁷. Proteins were fractionated by ammonium sulphate precipitation. Rad1 protein was further purified by DEAE-cellulose, heparin-Sepharose, FPLC Mono-Q and FPLC Superose-12 column chromatography. Using this purification scheme, about 100 μ g of apparently homogeneous Rad1 protein was obtained. It was estimated by immunoblotting that this represents a yield of $\sim$ 5%.

Escherichia coli or yeast (A.J.B. and E.C.F., unpublished observations) and after *in vitro* transcription and translation of the cloned *RAD1* gene¹⁶. Rad1 and Rad10 proteins translated *in vitro*, or present in yeast cell-free extracts, form a stable and specific complex^{16,17}, and a specific interaction between these proteins has also been demonstrated *in vivo*¹⁹, suggesting that Rad1 and Rad10 participate in repair and recombination as a protein complex.

When purified Rad1 and Rad10 proteins were incubated under conditions resulting in complex formation with circular single-stranded M13 or Φ X174 DNA, there was endonucleolytic degradation of the substrate. Figure 2 shows the progressive degradation of single-stranded DNA when increasing amounts of purified Rad10 protein were added to reaction mixtures in the presence of a fixed amount of purified Rad1 protein. This result was also obtained in the reciprocal experiment, in which increasing amounts of Rad1 protein were added to a fixed amount of Rad10 protein (data not shown). Neither protein alone degraded the substrate (Fig. 2). The Rad1/Rad10 endonuclease did not degrade linear DNA and no exonuclease activity was detected

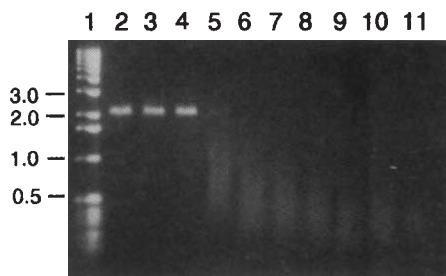


FIG. 2 Co-incubation of Rad1 and Rad10 proteins with single-stranded DNA results in endonucleolytic degradation. Lane 1, M_r standards (Gibco/BRL; 1-kb ladder). Reaction mixtures (20 μ l final volume, containing 50 mM HEPES-NaOH, pH 8.5, 20 mM NaCl, 1 mM DTT, 10 mM $MgCl_2$ (endonuclease buffer) plus 125 ng M13 single-stranded DNA) were incubated for 30 min at 37 $^{\circ}$ C with the following protein additions: lane 2, no protein added; lane 3, 1 pmol Rad1; lane 4, 2.5 pmol Rad10; lane 5, 1 pmol Rad1 and 0.25 pmol Rad10; lane 6, 1 pmol Rad1 and 0.5 pmol Rad10; lane 7, 1 pmol Rad1 and 0.75 pmol Rad10; lane 8, 1 pmol Rad1 and 1.0 pmol Rad10; lane 9, 1 pmol Rad1 and 1.5 pmol Rad10; lane 10, 1 pmol Rad1 and 2.0 pmol Rad10; lane 11, 1 pmol Rad1 and 2.5 pmol Rad10. Size markers (kb) are shown on the left. **METHODS.** Reactions were terminated by heating for 2 min at 100 $^{\circ}$ C to dissociate DNA-protein complexes. Degradation of the DNA was analysed by electrophoresis in native 1% agarose gel in the presence of ethidium bromide and visualized by ultraviolet transillumination.

using either duplex or single-stranded DNA (data not shown). The endonuclease activity has a requirement for Mg^{2+} , a pH optimum of 8.5, and is inhibited in the presence of NaCl at concentrations >20 mM.

The kinetics of the degradation of single-stranded DNA was further characterized using a quantitative assay. In reactions containing equimolar amounts of Rad1 and Rad10 proteins, endonucleolytic cleavage of M13 DNA was effected at an initial rate of 0.2 mol DNA ends formed per mol of Rad1-Rad10 protein per minute (Fig. 3). Once again, no degradation of the substrate was detected during extended incubation in the presence of an excess of either Rad1 or Rad10 protein alone (Fig. 3). Efficient labelling of free DNA ends produced by the endonucleolytic degradation required prior dephosphorylation, indicating that the Rad1/Rad10 cleavage products contain 5'P termini.

Rad10 protein binds preferentially to single-stranded DNA and promotes partial renaturation of complementary single strands¹⁸. We therefore considered the possibility that the protein may modify the structure of single-stranded DNA such that it becomes susceptible to subsequent endonucleolytic cleavage by Rad1 protein. But neither purified *E. coli* single-stranded binding protein nor bacteriophage T4 gene 32 protein, both of which bind to single-stranded DNA and promote strand annealing²⁰, could substitute for Rad10 protein in the endonuclease reaction (data not shown). These results suggest that a specific interaction between Rad1 and Rad10 proteins is required for endonuclease activity.

To confirm that Rad1/Rad10 protein complexes have endonuclease activity, we covalently attached purified Rad10 protein to Affi-gel beads. Rad1 protein bound stably to the Rad10-beads, but not to beads with covalently attached bovine serum albumin (BSA). This interaction was stable in 2M NaCl, but was sensitive to incubation in 1% SDS at 70 $^{\circ}$ C. Rad1/Rad10 protein complexes constituted on Affi-gel beads degraded single-stranded DNA (Fig. 4). Under these incubation conditions $<6\%$ of Rad1 protein dissociated from the Rad10-beads (data not shown). No endonuclease activity was detected with BSA-beads preincubated with Rad1 protein (Fig. 4), or with Rad10-beads alone (data not shown). Collectively our results indicate that Rad1/Rad10 protein complexes generated in the absence of DNA have single-stranded DNA endonuclease activity, and suggest that stable complex formation is required to constitute this activity.

These observations are compatible with the predicted roles of these gene products in both DNA repair¹ and intrachromosomal mitotic recombination¹³. In both yeast and mammalian cells, damage-specific recognition and incision of DNA requires

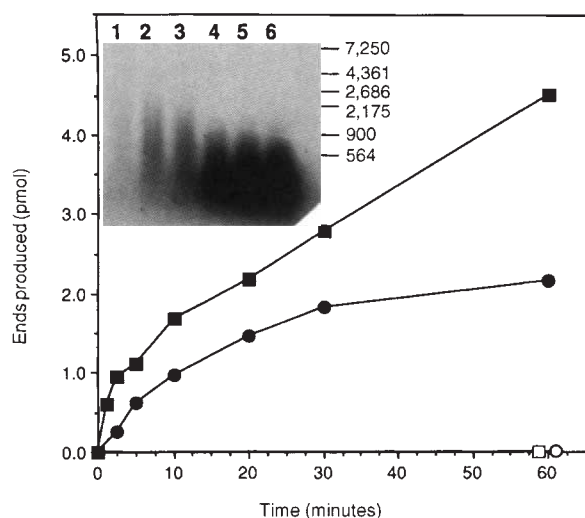


FIG. 3 DNA incision by Rad1 and Rad10 proteins as a function of time. Endonuclease reactions (20 μ l) contained 250 ng circular M13 single-stranded DNA and: 1 pmol each of Rad1 and Rad10 proteins (filled circles), 2 pmol each of Rad1 and Rad10 (filled squares), 10 pmol Rad1 alone (circles) or 10 pmol Rad10 alone (squares). Reactions were terminated at the indicated times and the DNA products were end-labelled. Labelled ends were quantitated by precipitation onto positively charged filters. The plotted points are the means of three separate experiments. Values were corrected for a labelling efficiency of 65%, determined in control experiments. Inset shows a denaturing agarose gel analysis of end-labelled products of reactions (20 μ l) containing 2 pmol each of Rad1 and Rad10. Lane 1, 2.5 min reaction; lane 2, 5 min; lane 3, 10 min; lane 4, 20 min; lane 5, 30 min; lane 6, 60 min. Sizes (in nucleotides) and positions of single-stranded molecular weight markers are indicated on the right. Measurement of the specific activity of the endonuclease based on the mass distribution of radiolabelled products was in agreement with that determined by quantitating DNA ends generated by endonucleolytic degradation.

METHODS. Endonuclease reactions were done as described for Fig. 2, except that Rad1 and Rad10 proteins were incubated for 30 min at 30 °C before adding substrate DNA. After incubation for the times indicated, 10- μ l aliquots were treated with calf intestinal alkaline phosphatase (Promega) and subsequently with [32 P]ATP and T4 polynucleotide kinase as described²⁸. Duplicate 5- μ l aliquots were then precipitated onto DE81 paper (Whatman) as described²⁸ and quantitated by liquid scintillation counting. Aliquots (15 μ l) were brought to 50 mM NaOH and electrophoresed on a 1% alkaline agarose gel and autoradiographed.

the products of several genes^{1,2}. One of the yeast genes (*RAD3*) encodes a DNA-helicase activity^{21,22} that is inhibited by base damage²³⁻²⁵. The yeast *SSL2* gene and its human homologue *ERCC3* also contain amino-acid sequence domains that are conserved in DNA helicases^{6,26}. Hence, localized unwinding of DNA mediated by one or more DNA helicases at or near sites of DNA damage may generate single-stranded regions that are cleaved by the Rad1/Rad10 endonuclease in yeast, and its functional homologue in mammalian cells. The Rad1/Rad10 endonuclease degrades single-stranded DNA at a relatively low rate, suggesting that the activity is modulated by other protein-protein interactions *in vivo*. Alternatively, the endonuclease may cleave a specific DNA structure(s) generated in living cells which includes single-stranded regions at or near sites of base damage.

The results of genetic studies on intrachromosomal recombination between directly repeated sequences, coupled with physical analysis of recombination intermediates, suggest that the *RAD1* gene product is required for productive recombination in the presence of 3' regions of non-homology that limit annealing and/or strand invasion¹³. The Rad1/Rad10 endonuclease may catalyse the removal of such non-homologous 3' single-stranded regions¹³. As in nucleotide excision repair, the Rad1/Rad10 endonuclease may exhibit increased activity in the presence of specific recombination intermediates and/or other proteins.

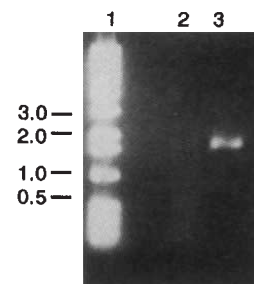


FIG. 4 The Rad1/Rad10 protein complex is a single-stranded DNA endonuclease. Rad1 protein (1.0 μ g) was incubated for 30 min at 30 °C with 50 μ l Rad10-beads or BSA-beads in immunoprecipitation buffer¹⁶, followed by extensive washing with this buffer to remove unbound Rad1 protein. Under these conditions ~0.5 μ g Rad1 protein was bound to the Rad10-beads. There was no detectable binding to the BSA-beads. Beads were then washed with endonuclease buffer (Fig. 2 legend) and resuspended in 20 μ l of the same buffer containing 250 ng M13 circular single-stranded DNA. Lane 1, *M*, standards; lane 2, Rad1 protein preincubated with Rad10-beads; lane 3, Rad1 protein preincubated with BSA-beads. After incubation for 60 min at 37 °C reactions were terminated as described in Fig. 2 legend. The beads were removed by centrifugation and DNA degradation was analysed as for Fig. 2. Size markers (in kb) are shown on the left.

METHODS. Rad10 protein was purified to homogeneity by a modification of the previously described purification schemes^{18,27}. About 2 mg of >95% homogeneous Rad10 protein was obtained with a recovery of ~20%. Rad10 protein (2 mg) and BSA (2 mg) were each covalently attached to 0.5 ml Affigel-10 beads (BioRad) according to ref. 29.

The human *ERCC1* gene is homologous to *RAD10* in the region of amino-acid identity known to be required for interaction with Rad1 protein^{3,16,19}. It is therefore likely that there is an as-yet unidentified mammalian gene that shares amino-acid sequence homology with the yeast *RAD1* gene, and that the product of this gene interacts with Erccl protein to generate a mammalian equivalent of the Rad1/Rad10 endonuclease. □

Received 17 December 1992; accepted 17 February 1993.

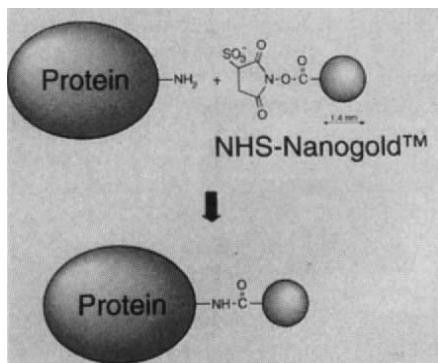
1. Friedberg, E. C., Siede, W. & Cooper, A. J. in *The Molecular and Cellular Biology of the Yeast Saccharomyces: I. Genome Dynamics, Protein Synthesis and Energetics* (ed. Broach, J., Jones, E. & Pringle, J.) 147-192 (Cold Spring Harbor Laboratory Press, New York, 1991).
2. Hoeijmakers, J. H. J. & Bootsma, D. *Cancer Cells* **2**, 311-320 (1990).
3. van Duin, M. et al. *Cell* **44**, 913-923 (1986).
4. Weber, C. A., Salazar, E. P., Stewart, S. A. & Thompson, L. H. *EMBO J.* **9**, 1437-1447 (1990).
5. Bankmann, M., Prakash, L. & Prakash, S. *Nature* **355**, 555-558 (1992).
6. Gulyas, K. D. & Donahue, T. F. *Cell* **69**, 1031-1042 (1992).
7. Friedberg, E. C. *Cell* **71**, 887-889 (1992).
8. Reynolds, P., Prakash, L. & Prakash, S. *Molec. cell Biol.* **7**, 1012-1021 (1987).
9. Reynolds, P., Prakash, L., Dumais, D., Perozzi, G. & Prakash, S. *EMBO J.* **4**, 3549-3552 (1985).
10. Aguilera, A. & Klein, H. L. *Genetics* **123**, 683-694 (1989).
11. Schiestl, R. H. & Prakash, S. *Molec. cell Biol.* **8**, 3619-3626 (1989).
12. Thomas, B. J. & Rothstein, R. *Genetics* **123**, 725-738 (1989).
13. Fishman-Lobell, J. & Haber, J. E. *Science* **258**, 480-484 (1992).
14. Schiestl, R. H. & Prakash, S. *Molec. cell Biol.* **10**, 2485-2491 (1990).
15. Petes, T. D., Malone, R. E. & Symington, L. S. in *The Molecular and Cellular Biology of the Yeast Saccharomyces: I. Genome Dynamics, Protein Synthesis, and Energetics* (ed. Broach, J., Jones, E. & Pringle, J.) 407-521 (Cold Spring Harbor Laboratory Press, New York, 1991).
16. Bardwell, L., Cooper, A. J. & Friedberg, E. C. *Molec. cell Biol.* **12**, 3041-3049 (1992).
17. Bailly, V., Sommers, C. H., Sung, P., Prakash, L. & Prakash, S. *Proc. natn. Acad. Sci. U.S.A.* **89**, 8273-8277 (1992).
18. Sung, P., Prakash, L. & Prakash, S. *Nature* **355**, 743-745 (1992).
19. Bardwell, A. J., Bardwell, L., Johnson, D. K. & Friedberg, E. C. *Molec. Microbiol.* (in the press).
20. Chase, J. W. & Williams, K. R. A. *Rev. Biochem.* **55**, 103-136 (1986).
21. Sung, P., Prakash, L., Matson, S. W. & Prakash, S. *Proc. natn. Acad. Sci. U.S.A.* **84**, 8951-8955 (1987).
22. Harosh, I., Naumovski, L. & Friedberg, E. C. *J. biol. Chem.* **264**, 20532-20539 (1989).
23. Naegeli, H., Bardwell, L. & Friedberg, E. C. *J. biol. Chem.* **267**, 392-398 (1992).
24. Naegeli, H., Bardwell, L. & Friedberg, E. C. *Biochemistry* **32**, 613-621 (1993).
25. Naegeli, H., Modrich, P. & Friedberg, E. C. *J. biol. Chem.* (in the press).
26. Weeda, G. et al. *Cell* **62**, 777-792 (1990).
27. Bardwell, L., Burtcher, H., Weiss, W. A., Nicolet, C. M. & Friedberg, E. C. *Biochemistry* **29**, 3119-3126 (1990).
28. Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* 2nd edn (Cold Spring Harbor Press, New York, 1989).
29. Formosa, T., Barry, J., Alberts, B. M. & Greenblatt, J. *Meth. Enzym.* **208**, 24-54 (1991).

ACKNOWLEDGEMENTS. These studies relied heavily on unpublished results (A.J.B.) which make an important contribution to our understanding of Rad1 protein. A.J.B. has previously published under the name A. J. Cooper. We thank our colleagues for discussion and D. Johnson for technical support. We also thank J. Fishman-Lobell and J. Haber for communicating before publication their suggestion concerning the possible endonucleolytic involvement of Rad1 protein. This work was supported by the USPHS.

Designer labels

A 1.4 nm gold compound, the biotinylated 21 contig probe, a kit for the purification of end-labelled probes and random-primed probes, enzyme immunoassays, antibodies, stains and catalogues are featured this week.

NANOPROBES has developed a new reagent, Mono-NHS-Nanogold, which is a 1.4 nm gold compound with a single organic reactive group, *N*-hydroxysuccinimide ester that reacts with primary amines under mild conditions (*Reader Service No. 100*). The



Mono-NHS-Nanogold gold compound reacts with primary amines.

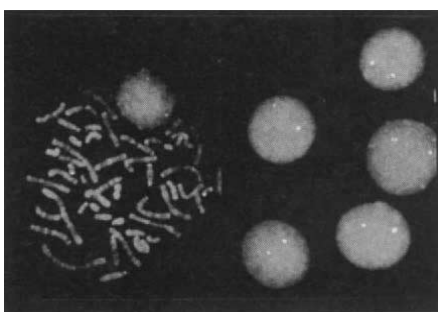
result, the company says, is a stable, covalent attachment of the Nanogold to a protein, lipid, peptide, DNA or other amine-containing target molecule. The reagent may be used to prepare gold probes for electron or light microscopy without the problems associated with colloidal gold (such as poorly controlled coupling, aggregation, poor stability and low diffusion into tissue). The gold can be enhanced with a silver developer (such as LI Silver or HQ Silver) so that the small gold particles are easily seen in the electron microscope, light microscope, or on blots and gels. It complements the company's existing product, mono-maleimido-Nanogold, which reacts with free thiols, and mono-amino-Nanogold, which is used in other reactions. Mono-NHS-Nanogold is designed to be easy to use. The dry mono-NHS-Nanogold is reconstituted with water and mixed with the target molecule (pH adjusted). After a few hours, the reaction is complete and the product can be purified by, for example, gel filtration, if desired. Nanoprobes supplies a full line of 1.4 nm gold covalently attached to antibodies and Fab' fragments, as well as streptavidin.

■ Hybridization

Amersham has introduced three fluorescent nucleotides and two new ³⁵S nucleotides for *in situ* hybridization (*Reader Service No. 101*). The new fluorescent nucleotides — fluorescein (FluoroGreen), rhodamine (FluoroRed) and coumarin (FluoroBlue) — can be incorporated into probes using nick translation or the polymerase chain reaction.

Following hybridization to chromosomes or nuclei, the probe can be detected directly using fluorescence microscopy. The use of different dyes allows multiple hybridizations on one preparation, Amersham says. FluoroBlue is not pH dependent. Stated applications include chromosome identification, whole chromosome mapping, typing karyotypes, translocation and single copy identification. [³⁵S]-UTPαS can be used to label the RNA probes for this application and has been developed to give both high specific activity and high concentrations to give increased sensitivity. [³⁵S]-dATPαS has been specifically formulated for use in 3'-end labelling of oligonucleotides for isotopic *in situ* hybridization. Amersham says that the new formulation offers extra stability throughout the half life of the product. It should be used with Amersham's 3'-end labelling kit.

The latest addition to Cambio's chromosome painting system, the **21 contig probe**, should be of interest to researchers studying Down's Syndrome (*Reader Service No. 102*). Developed at Cambridge University's



Biotinylated 21 contig probe.

Department of Pathology, the biotinylated 21 contig probe is prepared from two overlapping cosmids in the q22.3 region. Cambio says that it should be particularly useful for the detection of trisomy-21, and has been used successfully with uncultured amniotic fluid cells and cells in interphase. Each vial contains enough probe to process ten slides (26 × 32-mm coverslip). Separate kits are available for labelling and detection with either Texas Red or fluorescein isothiocyanate.

The new chromosome *in situ* **dual-colour detection system** from Alpha Laboratories enables researchers to use two different probes on the same sample, visualizing one as a red signal and one as green (*Reader Service No. 103*). This is designed to extend

the applications of *in situ* hybridization to: the identification and quantification of two or more chromosomes simultaneously in the same cell (the system is particularly suited for scoring X and Y chromosomes); to confirm reciprocal translocations (the *bcr/abl* translocation or the 'Philadelphia chromosome' can be visualized clearly); and to develop further colours (using other labelling systems researchers can extend the system to look at further probes). The kit is supplied with blocking, hybridization, detection and signal amplification reagents and is designed to provide maximum performance with any combination of Oncor's biotin- and digoxigenin-labelled probes, including alpha satellite, beta satellite, classical satellite, telomere, cosmid and coasome total chromosome probes.

A new kit for the rapid **purification of end-labelled probes and random-primed probes** has been developed by Advanced Genetic Technologies (*Reader Service No. 104*). A membrane that binds single-stranded DNA, but not mononucleotides, phosphate and biotin has been incorporated into a microcentrifuge cartridge. Labelled sample is loaded onto the cartridge and centrifuged for 5 s to remove unincorporated label. After another 5-s spin to wash the membrane, the company says that purified product with greater than 99 per cent purity is eluted in 80 µl TE. For random-primed probes, the sample is heat-denatured before purification on the cartridge, and then added directly to the hybridization mix. The company says that the DNA binding capacity is greater than 25 µg.

■ Assorted assays

AMAC's **soluble interleukin-2 receptor (sIL-2R) assay** is designed to be an easy one-step 'sandwich' enzyme immunoassay with a sensitivity of 0.5 pM (*Reader Service No. 105*). The standards, sample and conjugate are added to the monoclonal-



Soluble interleukin-2 receptor assay.

PRODUCT REVIEW

coated microtitre plate, a chromogenic substrate is added and incubated at room temperature for 30 min on a shaker. The plate can then be read on a spectrophotometer at 405 nm. AMAC supplies the sIL-2R enzyme immunoassay in a 96-test kit format containing all the reagents necessary to perform the assay.

An interleukin-7 'sandwich' enzyme immunoassay kit for the **quantitative determination of IL-7 levels** in human serum or cell culture supernatant is now available from Advanced Magnetics (*Reader Service No. 106*). Interleukin-7, also known as lymphopoietin 1, may act in regulating thymocyte growth and stimulating depleted



Interleukin-7 enzyme immunoassay kit.

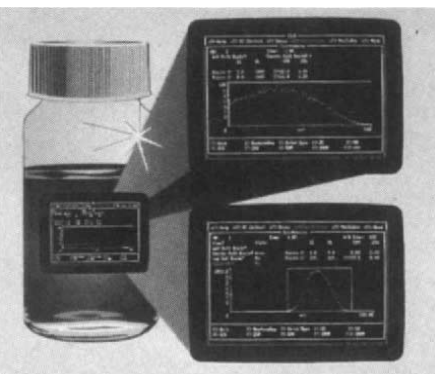
bone marrow. Precision-coated wells, supplied with each kit, are designed to provide consistency and flexibility. Prediluted, lyophilized standards eliminate variations in dilution, AMI says. The enzyme immunoassay is optimized for specificity, sensitivity and speed. The total assay time is 2 h 45 min, and the sensitivity in buffer is 4.1 pg ml⁻¹; 6.5 pg ml⁻¹ in serum. Interleukin-7 is specific for bioactive human IL-7. Sample preparation for this cytokine is not required. The kit contains sufficient reagents for 96 tests, including a standard curve.

The new **protein kinase assay kit** (PK Assay System) from Kamiya Biomedical offers a new approach to measuring the activity of protein kinase C and cAMP-dependent protein kinase (*Reader Service No. 107*). The system is based on an enzyme-linked immunosorbent assay that utilizes a synthetic peptide and a monoclonal antibody that recognizes the phosphorylated form of the peptide. The company says that it does not require the separation of phosphorylated and nonphosphorylated forms by agarose

gel electrophoresis, as in other systems. Kinase levels are measured colorimetrically at 450 nm. Results are available in less than 2 h, as compared to the 2 h 30 min to 3 h required by conventional methods, the company says. The PK Assay System is also less affected by ATP concentration in the sample. The ELISA is designed to offer the sensitivity of the radioactive method without the inconvenience and potential hazard. It can be run on crude cell extracts, column fractions or purified enzymes.

■ Reagents

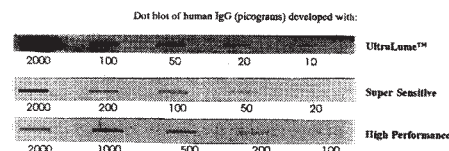
Ultima Gold AB is the new liquid scintillation cocktail from Packard that has been specially formulated for **alpha/beta discrimination** (*Reader Service No. 108*). It accepts a wide



Liquid scintillation cocktail formulated for alpha/beta discrimination.

variety of sample types for environmental monitoring or water, air, soil and ashed organic material. Packard says that it can handle aqueous sample loads of up to 50 per cent, even with high mineral acid concentrations, and can provide alpha/beta crossovers of well below 0.5 per cent. Ultima Gold AB is formulated from the diisopropyl naphthalene (DIN or DIPN) solvent, and is biodegradable. Its slow pulse decay characteristics make it suitable for alpha/beta discrimination with pulse decay analysis. It has a flash point of 150 °C, a low vapour pressure, good quench resistance up to its maximum sample capacity, and is virtually odourless, Packard says. It can also be diluted with the company's Ultima Gold F cocktail to optimize holding capacity for a wide range of samples and to reduce further alpha/beta crossover.

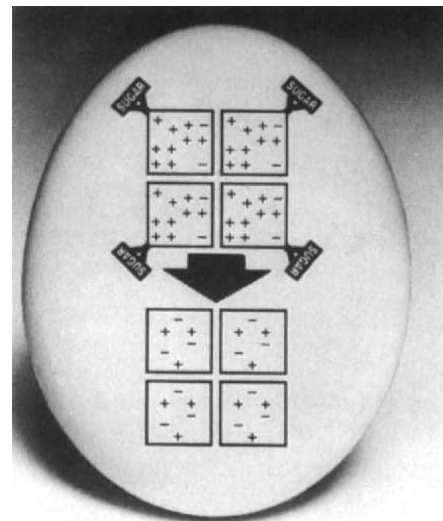
Two new chromogenic ImmunoBlot Detection Systems are available from BioGenex Laboratories in ready-to-use and concentrated formats for the **detection of human, mouse or rabbit antibodies** in western or dot blots (*Reader Service No. 109*). The high-performance system contains an enzyme conjugate that offers fast detection of 100 pg or less using a simplified protocol. The entire procedure can be completed in under 100 min, says BioGenex, and no immunoblotting experience or special skills are required. The Super Sensitive system



ImmunoBlot detection kits are designed to provide results in 100 min.

achieves enhanced sensitivity down to 20 pg or less with minimal background; this system utilizes an amplified biotin-streptavidin procedure that can be performed in about 2 h. The ImmunoBlot Detection Systems contain all the necessary detection reagents, including antibody diluent and wash buffer, in addition to step-by-step protocols. Each system is available in two different formats: convenient ready-to-use reagents (8-blot capacity) or as concentrated reagents (33-blot capacity). Both systems are compatible with the automated MAPS immunoblot processor.

NeutraLite Avidin from Accurate Chemical and Scientific is a protein-engineered product that is designed to offer a superior **alternative to streptavidin** in terms of cost and results (*Reader Service No.*



NeutraLite Avidin from Accurate.

110). The main advantage of NeutraLite Avidin is the absence of sugars, the result of enzymatically removed glycan. The product is neutral, free from nonspecific binding, with lysines still available.

Shandon OmniTags, which are **alkaline phosphatase detection kits** that form part of the Immunon range of immunocytochemical products, are now available from Life Sciences International (*Reader Service No. 111*). OmniTags kits use streptavidin-alkaline phosphatase as a label and contain sufficient reagents to stain up to 80 slides. They can be used with any primary antibody, such as rabbit, goat or mouse, through the use of a universal polyvalent biotinylated



OmniTags alkaline phosphatase detection kit (top) and CadenzaTags streptavidin-biotin immunochemical detection kit.

secondary antibody reagent. Antigens are stained red, contrasting with the blue haematoxylin counterstain. The advanced reagent alkaline phosphatase enhances the detection of antigens in routinely-processed, paraffin-embedded and frozen tissue sections, and blood or bone marrow films. OmniTag kits include a protein blocking agent, haematoxylin counterstain, chromogen tablets and equivalent buffer concentrate. The CadenzaTags kit is a three-step streptavidin-biotin immunochemical detection kit designed for the detection of ultrasensitive antibodies. The kits contain all the necessary reagents from the blocking step through to counterstaining (except primary), and can be complemented by Shandon's buffer concentrate and trypsin premixed with buffer salts. Based around a polyvalent secondary antibody reagent, the universal kit will detect rabbit, mouse, rat, guinea pig and sheep antibodies, reducing the need for laboratories to stock separate kits for different species. CadenzaTags are available with a choice of chromogen system. The kits are supplied with sufficient reagents to stain more than 300 slides.

Antibodies

Exalpa has released four **single-colour monoclonal antibodies** for use in flow cytometry and fluorescence microscopy: anti-CD35 (anti-CR1, complement C3b receptor), anti-CD11a (integrin LFA-1 α), anti-CD42a (anti-platelet) and a negative gamma 1-PE platelet control (*Reader Service No. 112*). The line of reagents can be used for the study of platelets, autoimmunity, T and B subsets, NK cells, T-cell activation, cellular adhesion and many other applications. Each vial is sufficient for 200 tests and each reagent is produced in a purified form or conjugated with fluorescein isothiocyanate, biotin or phycoerythrin. The company also offers a new CD3 PE monoclonal antibody that stains with high resolution (fourth quadrant

staining). Researchers can use Exalpa's CD3 PE for monitoring T-cell subsets, characterizing subtypes of T-cell leukaemias and lymphomas, studying AIDS/HIV infection, and in the analysis of CD3 complex related to the T-cell antigen receptor.

Novocastra Laboratories has developed a **mouse monoclonal antibody to oestrogen receptor** that can be used on Zamboni-fixed, frozen human tissue (*Reader Service No. 113*). It is now available from Vector Laboratories. Produced by using a synthetic peptide corresponding to a site of predicted high antigenicity on the human oestrogen receptor molecule as the antigen, this new clone provides a useful reagent for immunohistochemical research applications. When paired with the Vectastain Elite Mouse IgG Kit and DAB Substrate Kit from Vector Laboratories, this anti-oestrogen receptor antibody gives rise to an intense nuclear staining pattern in lobular and tubular breast carcinomas expressing oestrogen receptor without background staining. A mouse monoclonal antibody specific for human progesterone receptor protein has also been introduced and is available for immunohistochemical research. The anti-progesterone receptor antibody is produced through use of a synthetic peptide as antigen and can be used to stain Zamboni-fixed frozen human tissue. Its effectiveness on paraffin-embedded tissue is under investigation. The product is supplied as a lyophilized tissue culture supernatant sufficient to stain about 400 sections.

British Bio-technology has recently launched the Surfacemark range of products, which consists of purified and labelled **monoclonal antibodies for the analysis of cell surface antigens** (*Reader Service No. 114*). The range includes a variety of CD markers, reagents to detect cell proliferation, cell activation and isotype controls. The antibodies are offered in four formats: purified, biotinylated, and conjugated to fluorescein or phycoerythrin.

A **recombinant antibody (RC20) specific for phosphorylated tyrosine residues** is available from Affiniti Research Products (*Reader Service No. 115*). Developed by scientists at Transduction Laboratories and based on the monoclonal antibody produced by the clone PY20, this new molecule is a genetically engineered, antigen-binding fragment that has been directly conjugated to reporter molecules to facilitate ease of use. The company says that the complexes are sensitive and will not cross-react with phosphoserine or phosphothreonine residues. Used in conjunction with enhanced chemiluminescence technology, RC20 is capable of detecting phosphotyrosine-containing proteins as antibody concentrations down to 10 ng ml⁻¹ on western blots. As second-layer antibody conjugates are not required, the whole process from

epitope custom peptides

18 Windsor Terrace, Newcastle upon Tyne, UK, NE2 4HE.

HIGH QUALITY CUSTOM MADE PEPTIDES
COMPUTERISED EPITOPE PREDICTION
ANTIBODY PRODUCTION
PROTEIN SEQUENCING

PHONE OR FAX FOR A FREE QUOTATION

TEL: ++44 (0) 91 222 7186

FAX: ++44 (0) 91 221 0074

In association with Newcastle University Ventures

Reader Service No.24

ANTISENSE DNA/RNA

0.1 UM - 1000+ UM

S-Oligos DNA \$8/base

\$12/base large scale

O-Methyl RNA \$35/base

\$45/base large scale

800-888-2358; fax 800-869-0813

Intl.Fax 203-438-3637

Oligos Etc.

Reader Service No.16

\$3.50/Base

Custom DNA Synthesis

No Set-up Fee,
Shipped in 24 hours.

1-800-788-4362

1-800-786-4362 Fax

KEYSTONE
Laboratories, Inc.

Reader Service No.33

PACEMAKERSM

The multiple peptide synthesis and conjugation service

- High purity (up to 95%).
- Quantities for a range of uses.
- Single or multiple peptides.
- HPLC trace with each peptide.
- Terminal modification at no extra charge.
- Conjugation to carrier proteins.

from **£80 per peptide**

AFFINITI Research Products Ltd.
GPT Business Park, Technology Drive,
Nottingham, NG9 2ND, UK.
Tel: +44/0 602 436100; Fax: +44/0 602 436300

Reader Service No.23

PRODUCT REVIEW

initial blocking to final signal detection may be achieved in less than one hour. The product is available as horseradish peroxidase and alkaline phosphatase conjugates in two catalogue sizes.

Multiple Peptide Systems offers four **custom anti-peptide antibody services** to fit your research needs — standard gold, standard silver, fast gold and fast silver (*Reader Service No. 116*). The standard gold service is a complete antibody service that includes conjugation to a carrier protein, immunization of three rabbits, ELISA testing, approximately 150 ml of crude serum and affinity purification. The new standard silver service includes conjugation to a carrier protein, immunization of three rabbits, ELISA testing and approximately 150 ml of crude serum. Both services are available with a standard delivery schedule of 15–18 weeks. In addition, the fast gold or fast silver service is available with an accelerated delivery schedule of 10–11 weeks. All services are fully guaranteed and confidential.

O.E.M. Concepts has introduced **Sialyl Le^a specific monoclonal antibody**, which has been shown to be specific to the same tumour antigen that reacts with CA 19-9 (*Reader Service No. 117*). The immunogen used is human salivary mucin and specificity has been compared using purified antigen containing the CA-19 epitope and Sialyl Le^a polyacrylamide. Further testing on tissue samples by immunohistochemistry shows uniform reactivity with pancreatic cancer, colon cancer, stomach cancer and bile duct cancer. O.E.M. Concepts says that normal colon tissue is negative or positive only in outer parts of the crypts and in the glycocalyx. Normal pancreas tissue is positive only in the ductal tree. This antibody is an IgG₁ and has an affinity constant of 6×10^9 , which makes this antibody suitable for a variety of assay formats.

Options for hardware

SmartScan is a **PC-resident microplate reader** from Optometrics that will fit into the floppy disk drive slot of most PCs



PC-resident microplate reader.

(*Reader Service No. 118*). The device is also available as a standalone unit. Both configurations have a wavelength range of 340–800 nm. Standard 96-well microplates can be read by both models of SmartScan. The system is provided with a software package that controls SmartScan, acquires data and calculates and displays the results on the PC's monitor. A hard copy of the results can also be obtained.

CMA/Microdialysis has introduced the **CMA/280 fluorescence detector**, which can be used for liquid chromatography or flow injection analysis (*Reader Service No. 119*). Band filters for excitation and emission possesses maximal transmittance in the ranges 340–360 nm and 395–545 nm, respectively. These wavelengths are designed to cover the most common applications of fluorescence derivatization chemistries in HPLC. It is especially suitable for amino acids but is also suitable for compounds like histamine. The illuminated cell volume is 5.7 µl.

Stains

Fluorescent images can be obtained within minutes and without the toxicity of ethidium bromide with the new **rapid reversible visualization system for nucleic acids** on

nylon membranes, which can be obtained through NBS Biologicals (*Reader Service No. 120*). NBS says that the detection sensitivity is to 20 ng, when viewed with UV transillumination at 365 nm. The kit includes the destain reagent, which decolorizes the membrane for subsequent hybridization or other applications, and contains sufficient reagents to stain and destain 25 membranes using 40 ml per membrane.

It is often necessary to differentiate between glycoproteins and nonglycosylated proteins in a crude extract or enriched fraction, but direct identification of a glycoprotein can be difficult, especially if a faint glycoprotein band is to be targeted. In response to this problem, Boehringer Mannheim has developed the **DIG glycan/protein double labelling kit** that allows glycoproteins and proteins to be visualized in blue and reddish-brown, applying two simultaneous staining reactions on one track (*Reader Service No. 121*). Glycoproteins can be detected on western blots at concentrations of 10 ng and upwards; proteins from 50 ng upwards, Boehringer Mannheim says. The kit offers several additional options, such as the differentiation between glycoproteins containing sialic acids or not, the selective detection of terminal galactose residues of glycoproteins by using galactose oxidase, and the specific staining of particular sugar structures with DIG-labelled lectins.

Shelf stockers

The 28-page 1993–1994 **catalogue** from Olympus describes a complete line of monoclonal antibodies and related reagents for flow cytometry, cytokine and retrovirus research (*Reader Service No. 122*).

The Sigma Immunochemicals 1993 **catalogue** lists over 1,500 antibodies, growth factors, enzyme substrate and buffer tablets, and related products (*Reader Service No. 123*). □

These notes are compiled by Diane Gershon from information provided by the manufacturers. For more information, fill in the reader service card bound inside the journal.

ADVERTISEMENTS



Seymour Biotech Ltd.
CUSTOM SERVICES

- **OLIGONUCLEOTIDE SYNTHESIS**
High quality PROBES, PRIMERS, ANTISENSE OLIGO'S! Fast Efficient Service offering loads of modifications, BIOTIN, FLUORESCENCE, 5' & 3' labelling, Phosphothioates Affinity Support Oligos, etc.
- **MULTIPLE PEPTIDE SYNTHESIS**
Epitope Mapping & Screening! We now supply 5mg quantities of high purity custom peptides up to 15 residues in length, using Fmoc chemistry. With D or L amino acids, these are supplied lyophilised in 4ml vials, cleaved from solid phase & ready for use

Tel: 0562 825286 Fax: 0562 825284
Unit 23, Hookbrook Enterprise Centre
Worcester Road, Kidderminster DY10 1HY

Reader Service No.26

Synthetic Oligonucleotides
Rapid Quality Worldwide Service
£2.50 per base, no setup!
option of H.P.L.C. purification
biotinylation, fluorescent labelling
RNA Synthesis
phosphothioates, many other modifications
OSWEL DNA SERVICE
Department of Chemistry, University of Edinburgh
Kings Buildings, West Mains Road,
EDINBURGH EH9 3JJ
Tel: 031 650 4792 Fax: 031 662 4054

Reader Service No.18

Highest Quality Oligonucleotides
Synthesised, purified and shipped in 3 days
£2.75 per base
(All inclusive)
0.2 µM scale (approx. 500 µg)

- Antisense (completely or partially sulphurised)
- Extensive range of labelling and modifications
- Full design and analysis facilities
- Gene cassettes, bulk order and contract discounts

Molecular Medicine Unit
King's College School of Medicine & Dentistry
Tel. 071 - 326 3126 FAX 071 - 733 3877

Reader Service No.56

Japan opens up to foreign researchers

Opportunities for foreign scientists to gain employment in Japan have blossomed and can be expected to continue to grow in the coming years.

TEN years ago a foreign scientist was a novelty in a Japanese laboratory. Now it is rare to find a major laboratory without at least a few non-Japanese researchers, and many have several.

However, information on openings in the public and private sectors is not widely available. Most Japanese employers are unfamiliar with the classified advertising pages of international science journals, and those who are hesitate to use them. Instead, Japan expects foreign scientists to learn about such opportunities from their embassy officials in Tokyo or science organizations outside Japan, or through personal contacts with Japanese scientists (who often are equally ill-informed).

There is also little to help scientists considering a job in Japan. The govern-

ment has produced a few publications in English describing its research laboratories, for example, *National Laboratories and Research Public Corporations in Japan* published in 1990 by the Research Development Corporation Japan (JRDC) and *Daily Living in Japan*, written by the Agency of Industrial Science and Technology (AIST), but it spends nothing to publicise their existence.

In 1990, the Science and Technology Action Group (STAG) of the British Chamber of Commerce in Japan (BCCJ) published a guidebook for foreign scientists, *Gaijin Scientist*, in which some two dozen British scientists discuss working conditions in Japan (see *Nature* 349, 202; 1991). But the book is not easily available, and must be ordered direct from the BCCJ in Tokyo (see classified

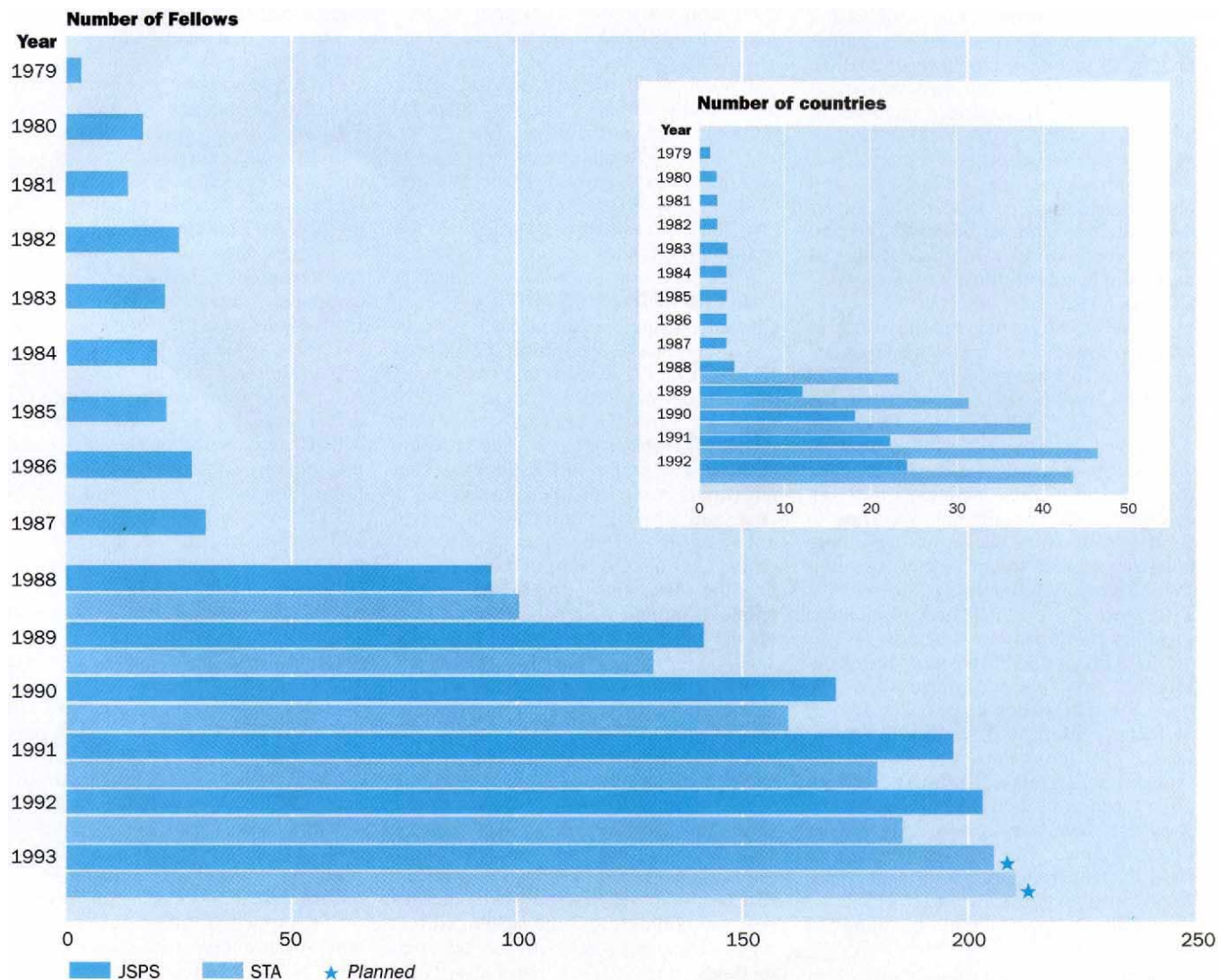
pages).

This special feature is intended to help close that information gap by providing a sense of the schemes and opportunities available to non-Japanese scientists.

Postdoctoral fellowships

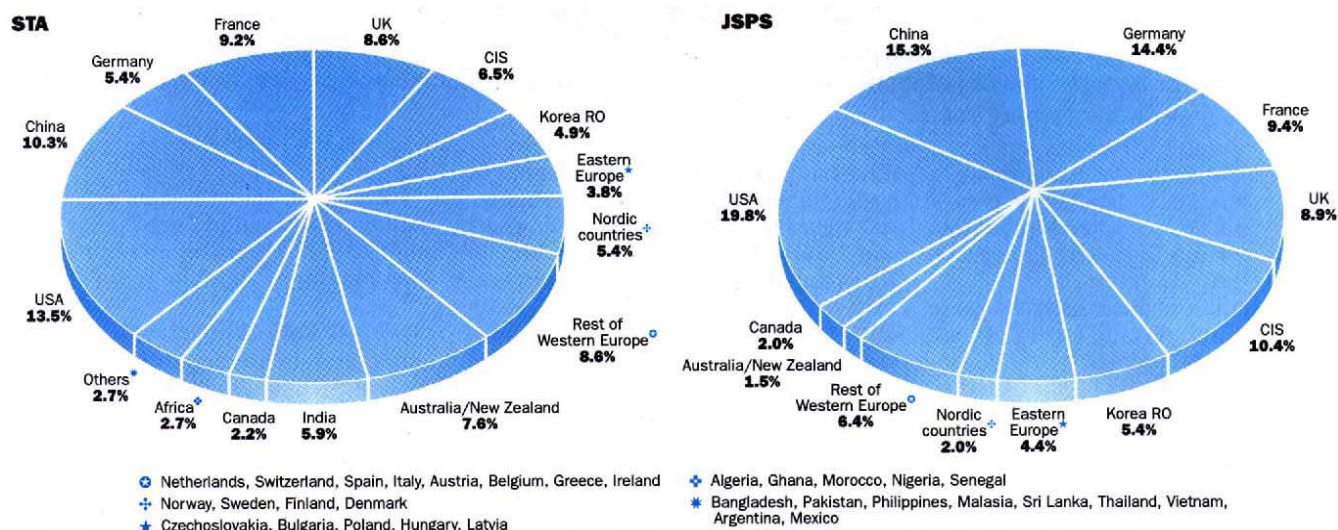
Japan began to open up its job market in the late 1980s in response to US criticism that the flow of Japanese scientists to the West far exceeded the influx of Western scientists to Japan.

In 1988, Japan's Science and Technology Agency (STA) began to offer postdoctoral fellowships for foreign scientists to carry out research at national laboratories (except universities), public corporations and non-profit research organizations. The Japan Society for Promotion of Science (JSPS), affiliated with the



EMPLOYMENT IN JAPAN

Share of Fellowships by country/region (1992)



Ministry of Education, Science and Culture (MESC), also expanded its programme for postdoctoral positions in universities beyond the three countries — United Kingdom, France and Germany — that had participated since 1979. These new opportunities coincide with a growing job market for foreign scientists in Japan's private sector (see page 870).

Japanese bureaucrats and academics are sceptical of the extent of the demand by foreign scientists to come to Japan, and at first that scepticism seemed justified (see *Nature* 335, 287; 1988). But once the fellowships were better advertised, the applications began to pour in. There are now two to three times as many applicants as positions under the STA and JSPS schemes.

In 1989, STA expanded the scope of its programme, and it now receives applications from scientists in more than 40 countries, including Eastern Europe, the former Soviet Union (CIS), India and Africa (see figures). JSPS has also increased the number of countries eligible for its postdoctoral scheme, but still excludes some developing countries in Africa and Asia. Scientists of these countries can, however, apply for other JSPS exchange programmes and in 1991 JSPS hosted over 1600 foreign scientists in Japan on its various schemes.

The STA and JSPS postdoctoral programmes now provide more than 400 one-year fellowships a year (see figure) with the possibility of extension to two years. The fellowships offer a tax-free monthly salary of ¥270,000 (US\$2,450) plus a generous housing allowance and Japanese language lessons as well as research funds. But fellows complain of rigid bureaucratic regulation of these funds, just as Japanese scientists complain about government handling of grants.

There are two ways that scientists may apply. Those in industrial Western na-

tions may apply through nominating organizations in their home country, such as the US National Science Foundation or the Royal Society in Britain (see footnote to table), or by filing applications through a host scientist or host institution. Scientists in non-Western countries must follow the second route.

In addition to the STA and JSPS programmes, there are several other postdoctoral schemes run by organizations in Japan and Europe (see table), and Japan's pharmaceutical industry has also introduced several small schemes (*Nature* 361, 762; 1993). In total well over 500 postdoctoral fellowships are available each year.

Top-rate laboratories

Choosing a suitable laboratory is difficult given the dearth of information about Japan. For STA fellows, one of the most popular destinations is the Institute of Physical and Chemical Research (RIKEN) outside Tokyo. The institute carries out research in the physical, chemical and biological sciences and has a large and growing community of foreign scientists, in part because of the STA scheme but also because of the institute's Frontier Research Programme which employs foreign scientists to lead some of the projects (see *Nature* 359, 578; 1992).

The nearly 50 national research laboratories in Tsukuba science city northeast of Tokyo also attract many foreign fellows. But applicants should take care to avoid institutes best described as 'dinosaurs'. The National Laboratory for High Energy Physics (KEK) affiliated with MESC (and thus open to JSPS applicants) is carrying out world-class research on its TRISTAN electron-positron collider, and the synchrotron at KEK's Photon Factory attracts scientists in fields ranging from material sciences and electronics to protein analysis.

The Electrotechnical Laboratory (ETL) affiliated to the Ministry of International Trade and Industry also attracts many fellows and has some scientists involved in international research, for example, under the Human Frontier Science Program.

The trade ministry has just reorganized four of its other institutes near ETL. Of particular interest to foreign scientists is the new National Institute for Advanced Interdisciplinary Research, which intends to hire outside researchers, including foreigners, to create a more flexible research environment. The institute will act as a centre for the new national project on nanotechnology involving several US companies, including Motorola, Texas Instruments and DuPont (see *Nature* 362, 279; 1993).

Elsewhere in Japan, the 'national institutes for joint university use', established since the 1970s and operated by MESC, have some of the best research facilities in Japan. Of note is the Institute of Space and Astronautical Sciences (ISAS) in Sagami-hara near Tokyo, world-renowned for X-ray astronomy, and the Nobeyama Observatory of the National Astronomical Observatory, which has some of the most powerful radio telescopes in the world. The three institutes (for basic biology, molecular science and physiological science) in Okazaki, near Nagoya, also have strong international reputations and a reasonable compliment of foreign scientists.

The Ocean Research Institute of Tokyo University and the Japan Marine Science and Technology Center have some of the world's most sophisticated research vessels and submersibles for marine research. And Japan has several well-equipped centres for earthquake and volcano research. But the persistence of the Japanese seismological community in maintaining that earthquake

prediction is possible places them behind developments elsewhere in the world (*Nature*, **356**, 464; 1993).

Several semiprivate institutes have opened in recent years that boast luxurious facilities, such as the Protein Engineering Research Institute in Osaka, the adjacent Osaka Bioscience Institute, the Advanced Telecommunications Research Institute International (ATR), also near Osaka, and the International Superconductivity Center (ISTEC) in Tokyo. And the Research Institute of Innovative Technology for the Earth (RITE), which will open a new facility this summer near ATR in the green surroundings of the Kansai science "city"

between Osaka, Nara and Kyoto, has already started recruiting foreign scientists (*Nature* **360**, 403; 1992).

Japan's national universities are generally cramped and in disrepair. The research environment, in which professors tend to yield considerable power over junior researchers, is often depressing, particularly for young foreign scientists. Nevertheless, many of Japan's best and most internationally minded researchers are in the leading national universities, such as Tokyo and Kyoto.

The names of such researchers can be found in top international journals and in lists of awardees of international grants, for example, the Human Frontier

Science Program. Also, in July MESC awards about a dozen large grants for "internationally distinguished research" (see, for example, *Nature* **358**, 446; 1993). English abstracts describing the projects are available from the Research Aid Division of MESC's International Affairs Bureau.

Permanent positions lacking

Although the number of postdoctoral positions for foreign scientists has increased dramatically in recent years, the opportunities for permanent employment in Japan's universities and national laboratories, remain limited (see *Nature* **345**, 380; 1990).

In 1982, Japan's laws were rewritten to allow national universities to hire foreign professors, associate professors and lecturers. The number of such foreign faculty in the 98 national universities now stands at about 200 (see table), although 85 per cent have temporary appointments lasting 1-5 years instead of tenure and fewer than half are working in science and engineering departments.

A recent external review of Tokyo University's physics department by foreign and Japanese scientists strongly recommended advertising all faculty openings on the international job market with no requirement that applicants speak Japanese (see *Nature* **362**, 387; 1993).

FOREIGN FACULTY STILL A MICRO MINORITY

Position	Foreign nationals		Japanese permanent
	temporary	permanent	
Professor	27	3	17,215
Associate professor	104	11	14,881
Lecturer	53	3	4,832
Total	184	17	36,928

In addition to above there are 63 faculty employed temporarily in endowed chairs. About half are non-Japanese.

But it seems unlikely that such a policy will become widespread in the near future; most Japanese faculty members are reluctant to employ foreigners, and administrators are not comfortable dealing with foreign faculty who do not speak Japanese.

Laws have also been changed to allow for long-term employment of foreign scientists in national institutes and permanent positions are easier to obtain than in the universities. But the numbers of such researchers are also still very small.

Applicants for postdoctoral positions who would like to be considered for full-time employment should question their prospective host scientist before coming to Japan. An alternative is to seek employment in Japan's private companies, which do offer long-term positions (see next page).

David Swinbanks

FELLOWSHIP SCHEMES FOR JAPAN

Organization	Location	Duration (years)	No.	Eligibility	Contact
Agency of Industrial Science and Technology (AIST), Ministry of International Trade and Industry (MITI)	AIST Institutes	0.5-1	20	PhD under 35, must have permanent post in home country	International R & D Cooperation Division, AIST 1-3-1 Kasumigaseki, Chiyoda-ku, Tokyo
Science and Technology Agency (STA)	All national institutes and public research corporations (except universities and university-affiliated institutes)	1-2	210	PhD, ideally under 35	Responsible organizations* or Research Development Corporation of Japan, 2-5-2 Nagato-cho, Chiyoda-ku, Tokyo
Japan Society for Promotion of Science (JSPS)	Universities, university-related institutes	1-2	205	PhD, ideally 25-38	Responsible organizations† or Exchange of Persons Division, JSPS, Yamato Bldg, 5-3-1 Kojimachi, Chiyoda-ku, Tokyo
Commission of European Communities	Any national research institute or university	1-1.5	25	EC nationals, good English PhD under 35	Commission of the European Communities (DEXII-B-3), Rue de Loi 200, B-1049 Brussels, Belgium
Foundation for Promotion of Cancer Research‡	Institutes or hospitals affiliated to Ministry of Health and Welfare	1-2	30-40	Cancer researchers	Foundation for Promotion of Cancer Research, 5-1-1 Tsukiji, Chuo-ku, Tokyo
Human Frontier Science Program (HFSP)	Any research institution	1-2	150§	Research on brain and biological functions	International HFSP Organization Tour Europe, 20 Place des Halles, 67080 Strasbourg Cedex, France
Matsumae International Foundation	Universities, institutes and private companies	0.5	20	Mainly for developing countries: PhD or MSc, under 40, no experience of research in Japan	Matsumae International Foundation, Shinjuku Tokai Bldg, 3-27-4 Shinjuku, Shinjuku-ku, Tokyo 160

* Australia: The Australian Academy of Science; Austria: Federal Ministry for Science and Research; Canada: Natural Science and Engineering Research Council; Germany: Alexander Von Humboldt-Stiftung; Finland: The Academy of Finland; France: Centre National De La Recherche Scientifique; Italy: Ministero Della Ricerca Scientifica E Tecnologica; Netherlands: Stichting Voor De Technische Wetenschappen; New Zealand: Ministry of Research, Science and Technology; Norway: The Royal Norwegian Council for Scientific and Industrial Research; Sweden: NUTEK; Switzerland: Swiss National Science Foundation; United Kingdom: The Royal Society; United States: National Science Foundation; European Communities: Commission of European Communities.

† For Australia, Austria, Canada, Finland, France, Germany, Switzerland, United Kingdom and United States the responsible organizations are as above (for STA fellowships); Belgium: National Foundation for Scientific Research; Italy: National Research Council of Italy; Netherlands: Netherlands Organization for Scientific Research; New Zealand: Department of Scientific and Industrial Research; Sweden: Royal Swedish Academy of Sciences.

‡ Part of 10-year Strategy for Cancer Control.

§ Nationals of any country can apply for a HFSP fellowship to Japan but the 150 fellowships are not just for Japan.

Doors open in the private sector

PARALLEL with the opening of the public sector there has been rapid growth in opportunities for foreign researchers in Japan's private companies. Between 1988 and 1991 the number of foreign researchers in the private sector more than tripled to 750 (see figure). Some large corporations now employ over a hundred foreign researchers each (see table, right). The recession has cut into recruitment. But recruitment can be expected to resume when Japan pulls out of recession in the next year or two.

Japanese companies anticipate a shortfall in the domestic supply of researchers towards the end of this century because of a rapidly declining student population and a tendency for students to turn their backs on careers in science. In a 1991 survey of companies by the Science and Technology Agency, nearly 50 per cent say they intend to increase their foreign researchers to over 1 per cent of their research workforce, well above current levels in even the most progressive companies, and about 6 per cent have targets over 10 per cent.

Leading the pack are electronics companies like Sony, NEC, Toshiba, and Hitachi, which have special schemes to attract foreign researchers (see table below). Sony has a particularly progressive policy and has employed a British researcher to help with its recruitment drive. Competition is keen, with only 4 out of 200 applicants succeeding in a recent recruitment campaign. Sony's long-term goal is to increase the number of foreign researchers to 10 per cent, over ten times the present level.

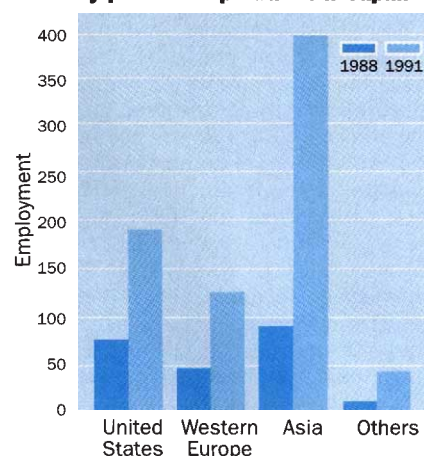
In the boom years of the late 1980s, many Japanese companies spent lavishly on new laboratories for basic research. These laboratories provide an attractive environment for researchers interested in long-term scientific research. An example is NEC's fundamental research

laboratory in Tsukuba science city which has 12 foreigners among the 300 research staff. The international teams of researchers there, for example, have produced a string of leading-edge research papers on carbon 60 and nanotubes (see *Nature* 362, 522; 1993).

Salaries for postdoctoral-level foreign researchers range from about ¥4 million (\$36,000) a year at the bottom of the scale to ¥9 million (\$81,000) a year for the very generous scheme run by Toshiba for British researchers. Most companies offer some sort of housing allowance on top of this.

Foreigners are usually employed on an

Foreign R&D personnel employed by private corporations in Japan



GROWING NUMBERS OF FOREIGN RESEARCHERS/ENGINEERS IN JAPANESE COMPANIES

Company	Foreign researchers/engineers						Researchers/engineers	Foreign employees	Employees
	1988	1989	1990	1991	1992	1993			
NEC	17	25	59	69	93	107	13,929	135	42,036
Toshiba	10	15	30	70	105	105	26,600	110	73,000
Sony	—	—	—	—	—	80	10,000	140	20,000
Mitsubishi Electric	—	—	—	—	54	80	23,200	68*	49,566
Fujitsu	—	—	—	—	—	~60†	—	98	54,442
Hitachi	37	54	53	41	40	40†	13,200	—	84,870
NTT	—	—	—	—	—	>30	8,600	—	242,700
Toyota	1	2	8	18	28	28§	14,250	59	73,000
Matsushita Electric	—	—	—	—	—	—	—	50	47,643
Okai Electric	3	6	17	20	28	30	—	36	14,050
Nissan	—	—	11	17	21	21	7,700	31	53,000
Canon	—	—	9	8	14	10	6,000	17	20,500
Kobe Steel	5	7	7	8	7	7	—	37	20,635
Sanyo	0	0	1	4	7	7	—	29	30,725
Mazda	—	—	—	—	—	1	5,000	17	26,621
Honda	—	—	—	—	—	0	11,400	10	45,000

*June 1992. †In addition, Fujitsu has about 30 visiting foreign researchers. ‡ Foreign researchers under Hitachi's HIVIPS visiting researcher programme. In addition to these, Hitachi currently has 5 permanent foreign researchers. Dashes indicate data not available throughout table.

§In addition, Toyota has about 100 foreign engineers temporarily assigned from overseas operations.

||Honda has about 100 foreign engineers temporarily assigned for 2–3 years from foreign affiliates.

annual contract basis. But some companies do offer the possibility of long-term employment on the same terms as Japanese colleagues. NEC, for example, has 8 such 'permanent' employees among 107 foreign researchers/

engineers. Sony encourages its foreign researchers to take up permanent employment after three years.

Japanese companies are recruiting researchers from many parts of the world. About half of NEC's foreign researchers come from advanced Western nations. The majority at Toshiba, on the other hand, come from Asia.

Japanese companies do not advertise positions widely, relying more on personal contacts. But many foreign researchers have succeeded in getting jobs simply by writing to companies.

While a few companies have well-developed policies for dealing with foreign researchers — Sony, for example, provides lessons in Japanese, and has in-house newsletters and videos in English to help new recruits settle in — others have a lot to learn. Those seeking employment in a Japanese company are advised to choose carefully. Companies with large numbers of foreign employees tend to have the most advanced policies on foreign recruitment.

D. S.

COMPANY SCHEMES FOR RESEARCH IN JAPAN

Company	Name of scheme	Eligibility	Contact
Hitachi	HIVIPS Visiting Researcher Programme	Professors or equivalent (level 1) PhD (level 2)	R & D Administration Office, Hitachi Central Research Laboratory, PO Box 2 Kokubunji, Tokyo
NEC	Visiting Researcher	Postdoctoral	Human Resources Development Division, International Personnel Relations, NEC Corp., 5-7-1, Shiba, Minato-ku, Tokyo
Sony	Global Researcher and Engineer Programme	University graduates	General Manager Human Resources, Sony Europa GmbH, Hugo-Eckener-Strasse 20, D-5000 Cologne, Germany
Toshiba	UK Fellowship Programme	Researchers in UK academic and government laboratories, ideally under 35	Toshiba Corporation (Europe Office), Audrey House, Ely Place, London EC1N 6SN, UK

Source: adapted from Gaijin Scientist (published by British Chamber of Commerce in Japan)